

Essays on Human Capital, Taxation, and Adverse Selection

by

Adam Blandin

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2016 by the
Graduate Supervisory Committee:

Gustavo Ventura, Chair
Todd Schoellman
Matthew Wiswall
Alexander Bick

ARIZONA STATE UNIVERSITY

May 2016

ABSTRACT

This dissertation consists of three chapters. The first two explore the impact of government policies on human capital accumulation. Chapter one makes two novel contributions related to the two workhorse models in the human capital literature: Learning by Doing (LBD) and Ben-Porath (BP). First, I show that BP is much more consistent with empirical life-cycle patterns related to individual earnings *growth rates* relative to LBD. Second, I show that the same model features that generate different life-cycle predictions between models also generate different policy implications. In particular, increasing the top marginal labor tax rate, relative to the current US level, generates much larger reductions in lifetime human capital accumulation in the BP model versus the LBD model.

Chapter two examines reforms to the Social Security taxable earnings cap in the context of a human capital model. Old age Social Security benefits in the US are funded by a 10.6% payroll tax up to a cap of \$118,500. There has been little work examining the likely outcomes of such a policy change. I use a life-cycle BP human capital model with heterogeneous individuals to investigate the aggregate and distributional steady state impacts of several policy changes the earnings cap. I find that when I eliminate the cap: (1) aggregate output and consumption fall substantially; (2) the role of endogenous human capital is first order; (3) total federal tax revenues are lower or roughly unchanged; (4) about 1/3 of workers are made worse off.

The final chapter studies the existence and optimality of equilibria in the presence of asymmetric information. I develop an equilibrium concept which corresponds to the presence of mutual insurance organizations for a class of adverse selection economies which includes the Spence (1973) signaling and Rothschild-Stiglitz (1976) insurance

environments. The defining features of a mutual insurance organization are that policy holders are also the owners of the organization, and that the organization can write policies for which the terms depend on the experience of the mutual members. In general the equilibrium exists and is weakly Pareto optimal. Further, all equilibria have the same individual type utility vector.

To my mom and dad, for always being my biggest supporters.

To Jack, Elly, and Lindsey, for continuing to be my best friends even from far away.

ACKNOWLEDGMENTS

This research has benefited greatly from the input of my advisors: Gustavo Ventura, Todd Schoellman, Alex Bick, and Matt Wiswall. Gustavo’s energy and vision ensured that I was always improving and moving forward. Todd’s ability to distill my initial ideas into simple and concrete statements has been invaluable, and has made me a much better thinker, writer, and presenter. Alex has been both my most patient sounding board and my toughest critic—both roles have been invaluable. Matt provided a different perspective to the group, which helped to answer many of my questions and avoid many wrong turns.

I would also like to especially thank Edward C. Prescott. Ed graciously invited me to spend two summers working with him at the Minneapolis Fed, and hired me as his research assistant for three of my years at ASU. His Center for the Advancement of Economic Efficiency organized several conferences which I benefited from while at ASU, and helped fund my travel when I was accepted to outside conferences. Most important, Ed taught me the value of clear language, careful measurement, and the link between the two. Thinking “What would Ed say?” has now become a standard step in my research process, and I am better for it.

I am indebted to many other ASU faculty members for their time and advice, including Ed Schlee, Natalia Kovrijnykh, Eleanor Dillon, Domenico Ferarro, Berthold Herrendorf, B. Ravikumar, Greg Veramendi, Manjira Datta, and Bart Hobijn. I also thank David Lagakos, who helped get this project off the ground during his time at ASU.

For many helpful conversations, I thank my ASU classmates, particularly Chris

Herrington, Kevin Donovan, Xican Xi, Rishabh Sinha, and Francesco Agostinelli.

I am also grateful for support from several organizations outside for ASU, including economists at the Minneapolis Fed (especially Victor Rios-Rull, Elena Pastorino, and Kyle Herkenhoff) and the St. Louis Fed (especially David Wiczer, Alejandro Badel, Chris Waller, Guillaume Vandenbroucke, and Carlos Garriga).

For sparking and fueling my interest in economics I thank my professors from Washington University in St. Louis, especially John Nachbar, Gaetano Antinolfi and Lee Benham.

Finally, the third chapter of this dissertation was co-written with Ed Prescott and John H. Boyd. It was a privilege to work with such talented scientists.

TABLE OF CONTENTS

| | Page |
|--|------|
| LIST OF TABLES | x |
| LIST OF FIGURES | xi |
| CHAPTER | |
| 1 LEARNING BY DOING AND BEN-PORATH: DIFFERENT LIFE-CYCLE PREDICTIONS AND POLICY IMPLICA- TIONS | 1 |
| 1.1 Introduction | 1 |
| 1.2 The Life-cycle Human Capital Models | 6 |
| 1.2.1 The Environment and the Worker's Utility Maximization Problem | 7 |
| 1.2.2 The Human Capital Production Functions | 9 |
| 1.2.3 Policy Implications of the BP and LBD Models | 11 |
| 1.2.4 Life-cycle Predictions in the BP and LBD Models | 12 |
| 1.3 Life-cycle Facts for US Workers | 13 |
| 1.3.1 Life-cycle Profiles Targeted During Calibration of LBD and BP Models | 15 |
| 1.3.2 Untargeted Life-cycle Profiles | 17 |
| 1.4 Setting Model Parameters | 20 |
| 1.4.1 Parameters Set Exogenously | 21 |
| 1.4.2 Parameters Targeted to Data: No Age-Variation in Param- eters | 22 |
| 1.4.3 Parameters Targeted to Data: Allowing Age-Variation in Parameters | 28 |
| 1.5 Earnings Growth in Ben-Porath and Learning by Doing | 31 |

| CHAPTER | Page |
|---|------|
| 1.6 The Impact of Tax Progressivity on Human Capital Accumulation for Ben-Porath and Learning by Doing | 35 |
| 1.6.1 Increasing Marginal Tax Rates on the Top 1% of Earners ... | 36 |
| 1.6.2 Change from the Current US tax scheme to a Proportional Scheme..... | 39 |
| 1.7 Conclusion | 42 |
| 2 REFORMING SOCIAL SECURITY: | |
| THE EARNINGS CAP AND ENDOGENOUS HUMAN CAPITAL | 44 |
| 2.1 Introduction..... | 44 |
| 2.2 A Simple Model of Social Security | 50 |
| 2.2.1 The Empact of the Earnings Cap on Labor | 50 |
| 2.2.2 The Impact of the Earnings Cap on Human Capital Invest- ment | 55 |
| 2.3 The Model Economy | 59 |
| 2.3.1 Demographics and Endowments..... | 59 |
| 2.3.2 Preferences | 60 |
| 2.3.3 Human Capital Accumulation..... | 60 |
| 2.3.4 Technology | 60 |
| 2.3.5 Income Taxes | 61 |
| 2.3.6 Social Security | 61 |
| 2.3.7 The Government Dudget Constraint..... | 62 |
| 2.3.8 The Decision Problem of an Individual | 62 |
| 2.3.9 Stationary Equilibrium | 64 |
| 2.4 Calibration of the Benchmark Economy | 66 |

| CHAPTER | Page |
|---------|--|
| 2.4.1 | Parameters Not Related to the Social Security Program 66 |
| 2.4.2 | Social Security Parameters 69 |
| 2.5 | The Impact of Reforming Social Security 70 |
| 2.5.1 | Reform 1: Eliminating the Earnings Cap and Reducing the Social Security Funding Deficit..... 72 |
| 2.5.2 | Reform 2: Eliminating the Earnings Cap and Reducing the Payroll Rax Rate 74 |
| 2.5.3 | Reform 3: Eliminating the Earnings Cap and Increasing Old Age Benefits..... 75 |
| 2.5.4 | Reform 4: Reducing the Earnings Cap..... 76 |
| 2.6 | Conclusion 79 |
| 3 | EQUILIBRIUM WITH MUTUAL ORGANIZATIONS IN ECONOMIES WITH ADVERSE SELECTION..... 81 |
| 3.1 | Introduction..... 81 |
| 3.2 | The Relationship to Other Equilibrium Concepts..... 83 |
| 3.3 | The Formal Analysis 85 |
| 3.3.1 | Working in the Utility Space..... 86 |
| 3.3.2 | Some Results 87 |
| 3.4 | Application to the Adverse Selection Insurance Environment..... 88 |
| 3.4.1 | The Rothschild and Stiglitz Two-Type Adverse-Selection Environment 89 |
| 3.4.2 | Relation to the Rothschild and Stiglitz Equilibrium..... 92 |
| 3.4.3 | The General Rothschild and Stiglitz Case..... 93 |
| 3.5 | Application to the Spence Signaling Environment 93 |

| CHAPTER | Page |
|--|------|
| 3.5.1 Relation to Spence's Valuation Equilibrium | 96 |
| 3.6 Conclusion | 96 |
| REFERENCES | 97 |
| APPENDIX | |
| A THE HYBRID HUMAN CAPITAL MODEL | 100 |
| B VARIANCE OF EARNINGS, WAGES, AND HOURS WORKED | 104 |
| C ROBUSTNESS FOR THE LIFE-CYCLE VARIANCE OF EARNINGS GROWTH | 106 |

LIST OF TABLES

| Table | | Page |
|-------|--|------|
| 1.1 | Parameters Set Exogenously..... | 21 |
| 1.2 | Parameters Targeted to Data Moments, No Age-Variation in Parameters. | 25 |
| 1.3 | Parameters Targeted to Data Moments, with Age-Variation in Parameters. | 29 |
| 1.4 | Steady State Impact of the Tax Change on the Top 1% of Earners. | 39 |
| 1.5 | Steady State Impact of the Tax Change on Aggregate Variables. | 41 |
| 2.1 | Model Parameters. | 71 |
| 2.2 | Steady State Impact of Reforms on Economic Aggregates. | 72 |
| 2.3 | Share of Aggregate Changes Attributable to High-Endowment Workers. | 74 |
| 2.4 | Steady State Impact of Reforms on Government Revenues and Expenditures. | 75 |
| 2.5 | Steady State Impact of Reforms on Welfare of Newborn Workers. | 76 |

LIST OF FIGURES

| Figure | Page |
|---|------|
| 1.1 Means for Earnings, Hourly Wages, and Hours Worked..... | 16 |
| 1.2 Life-cycle Variance of log Earnings in the Data. | 17 |
| 1.3 Variance of Wage and Earnings Growth in the Data. | 18 |
| 1.4 Means for the Data and the Models without Age-Varying Parameters. . | 26 |
| 1.5 Means for the Data and the Models with Age-Varying Parameters. | 30 |
| 1.6 Variance of log Earnings for the Data and the Models with Age-Varying Parameters. | 31 |
| 1.7 Variance of 1-Period Earnings Growth Rates in the Models and the Data. | 33 |
| 1.8 Mean Time Spent Learning in Ben-Porath and Learning by Doing. | 34 |
| 1.9 Schedule of Earnings Taxes. | 37 |
| 1.10 Schedule of Earnings Taxes. | 40 |
| 2.1 Impact of a Tax Increase on High Earners..... | 53 |
| 3.1 $\hat{u}_L(u_H)$ in the Two-Type Rothschild and Stiglitz Environment for (a) Small λ_L ; and (b) Large λ_L | 92 |
| 3.2 The Utility Space in the Two-Type Spence Signaling Environment | 95 |
| A1 $H(\alpha)$ as a Function of Investment | 102 |
| B1 Life-cycle variance of log Earnings, Wages, and Hours in the Data | 105 |
| C1 Life-cycle Variance of 5-year Growth Rates | 107 |
| C2 Variance of 5-year Growth Rates by Age, Education | 108 |
| C3 Variance of 5-year Growth Rates by Experience, Education | 109 |
| C4 Life-cycle Variance of 5-year Growth Rates, $w_j/w_{j-5} \geq .9$ | 109 |
| C5 Life-cycle Variance of 5-year Smoothed Growth Rates, $w_j/w_{j-5} \geq .9$... | 110 |

Chapter 1

LEARNING BY DOING AND BEN-PORATH: DIFFERENT LIFE-CYCLE PREDICTIONS AND POLICY IMPLICATIONS

1.1 Introduction

Many government policies affect incentives to accumulate human capital for working-age adults. A growing literature studies the impact of these policies by running policy experiments within structural life-cycle models of human capital accumulation. The literature is split between two workhorse human capital models: Learning by Doing (LBD) and Ben-Porath (BP). In LBD workers learn automatically when they spend time producing; in BP learning and producing are separate, mutually exclusive activities.

This paper contains two novel contributions related to these models. First, I show that LBD and BP generate different life-cycle predictions. While both models are consistent with the life-cycle mean and variance of earnings, wages, and hours *levels* in the US, the BP model is much more consistent with life-cycle patterns related to individual earnings *growth rates*.

Second, I show that the same model features that generate different life-cycle predictions between the two models also generate different policy implications. The policy change I focus on in this paper is an increase in marginal labor tax rates for top earners. Increasing the top marginal tax rate reduces human capital accumulation by top earners to a greater extent in the BP model than in the LBD model. The end

result is that an increase in top marginal labor tax rates generates less revenue in the BP model relative to the LBD model.

My analysis is based on a life-cycle human capital model where workers are born as young adults endowed with an initial human capital stock and a learning ability, which determines their proficiency at accumulating future human capital. Heterogeneity in these initial endowments is the only way individuals differ from one another. Workers are also endowed with a unit of time in each period. Time can be split between three activities: production, during which the worker uses his human capital to produce goods and services; investment, during which no goods or services are produced; and leisure. In the BP version of the model workers learn by spending time investing in their human capital. In the LBD version workers learn by spending time in production, i.e. learning and earning occur simultaneously.

I first calibrate each model to the well-documented life-cycle profiles of mean earnings, mean hours worked, and the variance of earnings for employed men aged 23-60 in the Panel Study of Income Dynamics (PSID). The key features of these profiles are that the mean earnings profile is hump-shaped, the hours profile is roughly flat, and the variance of earnings profile is positively sloped.

I find that both models can quantitatively reproduce these profiles, but that they do so in different ways. In the BP model workers choose to invest less as they age (since the return to learning is lower for workers closer to retirement) and produce more as they age (since the return to producing, which is proportional to human capital, tends to be higher for older workers). Declining life-cycle investment endogenously generates a hump-shaped earnings profile because workers accumulate human

capital more slowly as they age. The BP model also generates a roughly flat mean hours profile because rising production time and falling investment time partially offset, resulting in small life-cycle changes to total hours worked.

Alternatively, workers in the LBD model cannot separately adjust time spent learning and earning. In order to simultaneously generate a hump-shaped earnings profile and a flat hours worked profile, then, the LBD model requires a combination of decreasing returns in the human capital function, as well as large life-cycle declines in learning ability or life-cycle increases in the rate of human capital depreciation.

In other words, both models predict that human capital (and earnings) grows more slowly for older workers, but for different reasons. With LBD, accumulating human capital simply becomes harder with age. With BP, on the other hand, workers *choose* to spend less time learning as they age.

These different predictions—that time spent learning falls to zero over the life-cycle in the BP model, but is roughly constant over the life-cycle in the LBD model—leads to different predictions for the distribution of earnings growth. BP predicts (1) early in the life-cycle, when workers are investing, the earnings of high ability workers grow faster than the earnings of low ability workers, and (2) later in the life-cycle, when workers have stopped investing, the earnings of all workers grow at similar rates. By contrast, because workers learn automatically in the LBD model, LBD predicts that the earnings of high ability workers grow relatively faster than the earnings of low ability workers throughout the life-cycle. Therefore, BP predicts that the variance in earnings growth rates should fall over the life-cycle, while LBD predicts that this variance should remain roughly constant over the life-cycle. In the data the variance

of 5-year earnings growth rates falls by 30-50% from age 30-55; this decline is robust to controlling for a host of factors including education, occupation, and temporary earnings shocks. The BP model accounts for 90% of the observed life-cycle decline in the variance of earnings growth, while the LBD model accounts for less than a third.

The different life-cycle predictions of these models also point to different policy implications. In both models, the return to learning falls to zero over the life-cycle of a worker. Workers in the BP model respond by dramatically reducing their human capital investment as they age. Workers in the LBD model hold the amount of time they spend learning roughly constant as they age. This suggests that a *policy* that lowers the return to learning will generate a larger response by workers in the BP model.

To illustrate this quantitatively, I analyze the steady state impact of increasing the marginal labor tax rate for the top 1% of workers in each model. I find that for a given tax increase, life-cycle time spent learning among workers affected by the tax increase falls much more in the BP model than in the LBD model, which causes life-cycle human capital to decline much more under the BP model. In particular, when I increase the marginal tax rate on the top 1% of earners to 73%, which is roughly the rate several previous papers have argued will maximize revenue from top earners, I find that time spent accumulating human capital among affected workers falls by 44% in the BP model versus 16% in the LBD model. This causes the earnings of these workers to fall more in the BP model, and the end result is that tax revenues from these workers falls by 36% in the BP model versus only 7% in the LBD model.

Background Many papers have examined the life-cycle predictions of either LBD or BP in isolation.¹ Only a pair of papers study the life-cycle predictions of LBD and BP side by side. Wallenius (2011) examines the predictions of a LBD and BP model for mean hours and wages over the life-cycle, paying special attention to the role of the intertemporal elasticity of substitution of labor. Fan *et al.* (2015) ask whether life-cycle human capital models are consistent with the behavior of wages and labor force participation in a model with endogenous retirement. Neither find compelling differences in the life-cycle predictions of LBD and BP. This paper is the first to show that LBD and BP generate different life-cycle predictions for the dispersion in earnings growth rates.

Two papers have pointed out specific policies whose implications differ across LBD and BP. Heckman *et al.* (2002) demonstrate that wage subsidies discourage skill formation if a BP technology is assumed, but increase skill formation in a LBD setting. Peterman (2012) demonstrates that the optimal tax rate on capital is 35% higher for a BP model than for a LBD model. A related project by Hansen and İmrohoroğlu (2009) demonstrates that the predicted volatility of hours worked over the business cycle differs across LBD and BP. I contribute to this literature by demonstrating an additional policy where the impact differs strongly across LBD and BP, and connecting the different policy implications of these models to different life-cycle predictions.

Finally, the tax policy analysis I conduct is relevant to a large literature on the impact of raising marginal tax rates on high earners.² It is most closely related to

¹Important examples include Shaw (1989) and Imai and Keane (2004) for the LBD model and Ben-Porath (1967), Heckman (1976), and Huggett *et al.* (2011) for the BP model.

²See for example Diamond and Saez (2011); Kindermann and Krueger (2014); Brüggemann and Yoo (2015); and Guner *et al.* (2015).

recent work by [Guvenen *et al.* \(2014\)](#) and [Badel and Huggett \(2014\)](#), which demonstrate that disproportionately raising marginal tax rates on high earners discourages human capital accumulation. Guvenen and coauthors use this idea to argue that differences in the progressivity of income taxes partially explain why (before-tax) wage inequality is larger in the US than in Europe. Badel and Huggett apply this insight to show that the revenue-maximizing marginal tax rate on top earners in the US is much lower in a model with endogenous human capital than in a model with exogenous wages. Both papers assume a BP human capital model. My policy experiments suggest that their assumption is quantitatively important: earnings and time spent learning respond much more to a change in tax progressivity in a BP model relative to a LBD model. Additionally, my finding that the life-cycle predictions of the BP model are consistent with a larger number of empirical moments than the LBD model provides support for their choice of a human capital model.

The remainder of the paper is organized as follows. Section 1.2 presents the LBD and BP versions of the life-cycle model. Section 1.3 documents life-cycle facts for US workers which I use to discipline the human capital models. Section 1.4 calibrates parameters for the LBD and BP models to match facts from Section 1.3. Section 1.5 then compares the LBD and BP model predictions for the variance of earnings growth rates over the life-cycle, which is untargeted during calibration. In Section 1.6 I analyze the impact of increasing the top marginal labor tax rate within each model, and Section 1.7 concludes.

1.2 The Life-cycle Human Capital Models

This section lays out the human capital accumulation model. I first discuss the economic environment and the lifetime utility maximization problem of an individual

in Section 1.2.1. Section 1.2.2 details the BP and LBD versions of the human capital production function. I then describe a worker's time allocation problem in each setting.

1.2.1 The Environment and the Worker's Utility Maximization Problem

Time is discrete. Every period a unit mass of workers are born. All workers live for \bar{J} periods. At birth each worker is endowed with initial human capital h_1 , and ability a which determines their proficiency in creating new human capital. The distribution of endowments (h_1, a) is given by $\Lambda(h_1, a)$; this is the only source of heterogeneity in the model.

Individuals are also endowed with a unit of time in each period. Time can be split between three activities: production, n , investing, s , and leisure, $1 - n - s$. The theoretical distinction between investing and production is that the individual is only paid a wage for time spent in production; time spent investing is useful only to the extent that it contributes to future human capital.

There is a single consumption good, c , whose price is normalized to one in each period. Individuals have time-separable preferences over consumption and leisure represented by a potentially age-specific utility function u_j . The individual's problem is to maximize the discounted value of lifetime utility subject to constraints on spending and human capital:

$$V(\bar{h}_1, a) = \max_{\{n_j, s_j, h_{j+1}, c_j\}_{j=1}^{\bar{J}}} \sum_{j=1}^{\bar{J}} \beta^{j-1} u_j(c_j, 1 - n_j - s_j) \quad (1.1)$$

$$s.t. \quad \sum_{j=1}^{\bar{J}} R^{1-j} c_j \leq \sum_{j=1}^J R^{1-j} \omega h_j n_j ; \quad (1.2)$$

$$h_{j+1} = (1 - \delta) h_j + H(h_j, n_j, s_j; a) ; \quad (1.3)$$

$$n_j, s_j \geq 0; \quad n_j + s_j \leq 1; \quad n_j = s_j = 0 \quad \forall j > J. \quad (1.4)$$

β is the individual's discount factor, and R is the gross interest rate in the economy. Constraint (1.2) states that the present value of lifetime consumption spending cannot exceed the present value of lifetime labor earnings. Labor earnings in period j are given by $\omega h_j n_j$, where ω is the wage rate for a unit of human capital; h_j is the individual's human capital in period j ; and n_j is the fraction of the time endowment spent producing in period j . $J \leq \bar{J}$ is an exogenous retirement date, *i.e.* individuals must set leisure equal to one beyond this date. From period $J+1$ to \bar{J} , the individual simply consumes the portion of his lifetime income which has not already been consumed.

The constraints in (1.3) describe the human capital accumulation process. Human capital in period $j+1$ is the sum of undepreciated human capital from the beginning of period j and human capital produced in period j . δ is the depreciation rate of existing human capital, and H is the human capital production function. The amount of newly produced human capital is a function of the individual's production and investing choices in period j , as well as his existing human capital and his learning ability.

1.2.2 The Human Capital Production Functions

There are two versions of the human capital production function H . The LBD version is given by:

$$H = ah^\phi n^\theta . \quad (1.5)$$

a is the individual's learning ability: individuals with higher ability produce more human capital for a given vector (h, n, s) than individuals with lower ability. ϕ and θ determine the curvature of new human capital with respect to current human capital and production time, respectively. This represents a pure LBD human capital accumulation process because the only input to human capital formation is time in production.³

The BP version of the human capital production function is given by:

$$H = ah^\phi s^\theta . \quad (1.6)$$

The function closely resembles the LBD case, except that now investment, s is the sole time input for human capital formation. In the special case $\phi = \theta$ the function in (1.6) corresponds to the original [Ben-Porath \(1967\)](#) specification.^{4 5}

Despite the apparent similarity of the BP and LBD human capital functions, the time allocation problem of a worker looks very different across these two cases. In the LBD case investment is useless, so workers set $s_j = 0 \forall j$. Workers then make their time allocation decision in each period by equalizing the marginal benefit of leisure

³For examples of LBD functions previously considered in the literature see [Shaw \(1989\)](#); [Chang et al. \(2002\)](#); [Imai and Keane \(2004\)](#); or [Gemici and Wiswall \(2014\)](#).

⁴The human capital function in [Ben-Porath \(1967\)](#) also includes market goods, e.g. books or tutors, as inputs to human capital formation. The function in (1.6) implicitly assumes that the weight on human capital inputs produced in the market is zero.

⁵Papers using this functional form include [Heckman et al. \(1998\)](#); and [Huggett et al. \(2011\)](#).

and production:

$$\underbrace{-\frac{\partial u_j}{\partial n_j}}_{\text{MV(Leisure)}} = \underbrace{\lambda \omega h_j}_{\text{MV(Earning)}} + \underbrace{\lambda \left(\frac{\partial h_{j+1}}{\partial n_j} \right) \left[\sum_{k=1}^{J-j} R^{-k} \omega n_{j+k} \left(\frac{\partial h_{j+k}}{\partial h_{j+1}} \right) \right]}_{\text{MV(Learning)}} . \quad (1.7)$$

The left hand side of (1.7) is the marginal value of leisure. The right hand side is the marginal value of production, where λ is the Lagrange multiplier for the budget constraint. The marginal value of production in the LBD model is made up of two components. The first component is the marginal value of the worker's current effective wage, ωh_j . The right-most component in (1.7) is the marginal value of learning in period j , which is the marginal increase in discounted earnings throughout the remainder of the worker's life due to a marginal increase in production in period j . This component is absent in standard life-cycle models with exogenous wages.

In the BP model, solving a worker's time allocation decision requires two first order conditions:

$$\underbrace{-\frac{\partial u_j}{\partial n_j}}_{\text{MV(Leisure)}} = \underbrace{\lambda \omega h_j}_{\text{MV(Earning)}} = \underbrace{\lambda \left(\frac{\partial h_{j+1}}{\partial s_j} \right) \left[\sum_{k=1}^{J-j} R^{-k} \omega n_{j+k} \left(\frac{\partial h_{j+k}}{\partial h_{j+1}} \right) \right]}_{\text{MV(Learning)}} . \quad (1.8)$$

The first equality sets the marginal value of leisure equal to the marginal value of production, which in the BP model is simply the marginal value of the worker's current wage. The second equality sets the marginal value of production equal to the marginal value of investment. This additional equality arises because workers in a BP world face a tradeoff between earnings today and earning potential tomorrow, in contrast to workers in a LBD world who earn and learn through the same activity, production.

Workers who face such a tradeoff will behave differently from workers who do not. When the returns to different activities change, either due to ageing or to a policy

change, workers in a BP world will separately adjust production and investment to bring their marginal benefits back into balance. Workers in a LBD world, on the other hand, cannot separately adjust the time they spend earning and learning. The remainder of this section discusses how this leads to different life-cycle predictions and policy implications across the BP and LBD models.

1.2.3 Policy Implications of the BP and LBD Models

A simple example of a policy whose impact differs across the BP and LBD models is a temporary wage subsidy (for a more detailed analysis, see [Heckman *et al.* \(2002\)](#)). Ignoring government budget requirements and general equilibrium effects, suppose that the government unexpectedly subsidizes worker wages in the period j at some rate $\Delta > 0$. If the subsidy has only a small impact on a worker's lifetime income, then the only change to the worker's problem would be that the $MV(Earning)$ terms in equations (1.7) and (1.8) would change from $\lambda\omega h_j$ to $\lambda\omega(1 + \Delta)h_j$.

The wage subsidy raises the return to earning in period j , and in both models workers reduce leisure in period j to restore the first order condition between leisure and production. However, in the BP model investment in period j also declines in order to restore the second equality in (1.8). Therefore, the wage subsidy *decreases* human capital creation in the BP model, and *increases* human capital creation in the LBD model. The intuition is that the subsidy decreased the marginal value of learning relative to the marginal value of earning, which caused time spent earning and learning move in opposite directions from each other in the BP model, but move in the same direction in the LBD model (since learning and earning both occur via production in this model).

In Section 1.6 I analyze another policy which changes the marginal value of learning relative to earning for top earners: increasing the marginal labor tax rates imposed on top earners. I verify that this policy change generates different implications between the BP and LBD models, and that the differences are quantitatively important.

The returns to learning and earning don't just vary across policy regimes—they also vary over the life-cycle of a worker. Combined with the above finding that workers in BP and LBD models respond differently to a change in the returns to learning and earning, this suggests two points. First, BP and LBD will generate different life-cycle predictions. Second, empirical life-cycle patterns will be informative about how workers respond to changes in the relative returns to learning and earning, and should therefore be informative about how workers will respond to policies which impact these relative returns.

1.2.4 Life-cycle Predictions in the BP and LBD Models

The returns to earning and learning change over the life-cycle. The marginal value of earning activities grows as the worker's human capital grows. At the same time, all else equal the marginal value of learning tends to fall as workers age since there are fewer remaining periods before retirement in which new human capital can be used. Therefore, to continue to satisfy equation (1.8), workers in the BP model reduce investment and increase time in production over the life-cycle. This means that in the BP model the majority of human capital creation occurs early in the life-cycle; later in the life-cycle workers simply cash in on their prior investments by spending time in production. As noted in [Ben-Porath \(1967\)](#), an implication of this is that the BP

model endogenously predicts that the growth of hourly wages and earnings tend to slow over the life-cycle.

In the LBD model the marginal values of wages and learning evolve over the life-cycle just as they do in the BP model. However, since production is responsible for both earning and learning, workers only respond to changes to the *sum* of the value of their wage and of learning. Since the value of the former tends to rise over the life-cycle and the value of the latter tends to fall, it is not obvious what the life-cycle path of production and wages will be in the LBD model.

The central objective of this paper is to investigate the extent to which each of these models is consistent with empirical patterns of annual earnings, hourly wages, and hours worked over the life-cycle. Section 1.3 documents these patterns for US males in the PSID. In Sections 1.4-1.5 I demonstrate that the BP model is quantitatively consistent with these empirical patterns. I also demonstrate that, even when I allow model parameters to vary with age, the LBD model is quantitatively consistent with some moments but is inconsistent with others.

1.3 Life-cycle Facts for US Workers

This section establishes the main life-cycle facts for US workers. Section 1.3.1 documents life-cycle profiles for the mean and variance of annual earnings, hourly wages, and hours worked, which I use to set model parameters for both the LBD and BP versions of the model. Section 1.3.2 documents life-cycle patterns related to the growth of earnings and hourly wages. These moments are not targeted during calibration.

I use data on earnings and hours from the 1968-2009 family-level files of the Panel Study of Income Dynamics (PSID). I restrict my sample to observations of male heads of households between 23-60 years of age who report working at least 500 hours in a given year and whose labor earnings were between \$5,000 and \$1 million in 2009 dollars.⁶ These qualifications result in a sample of 11,145 individuals and 80,427 year-individual observations.

I exclude males over 60 for two reasons. First, the PSID contains few observations of working men over 60. Second, decisions about wages and hours at this stage in life may be exceptionally influenced by factors which are not present in this model.

I restrict ages on the low end in an attempt to focus on “working age” human capital accumulation, as opposed to human capital accumulated during primarily “school age” years. Two comments are necessary here. First, this restriction does not imply that I ignore early-life human capital accumulation. Rather, I simply capture the *result* of this early accumulation in a reduced form manner, through the initial distribution of human capital and ability endowments. In fact, since these initial endowments are the only source of heterogeneity in the model, early human capital accumulation plays a crucial role in my analysis. Second, we know that young individuals work very little before their formal schooling is complete, and that individuals who attend school longer earn higher hourly wages in their first few years of employment. A LBD model of early human capital accumulation would struggle to generate

⁶The sources of labor income used were “wages and salary”, “commission”, “overtime”, “professional practice income”, “tips”, “additional income”, and “all other labor income”. The minimum hours and earnings thresholds are designed to exclude individuals who were full time students or who experienced a long term unemployment spell. I throw out extremely high earners because these few outliers occasionally alter the measured moments.

these basic facts, while a simple BP process would do so easily. In contrast to this trivial analysis, we know relatively little about which human capital theories are consistent with hours and earnings facts after formal education has been completed.

I now discuss how the age profiles plotted in this section are obtained. The human capital model laid out in Section 1.2 is not able to replicate every aspect of time variation contained in the data, so I use a statistical model to extract the sources of variation the model is designed to address. Following the approach of [Huggett *et al.* \(2011\)](#), I first group the data into 5-year centered age bins. I then assume that the statistics of interest are governed by the following fixed effects model:

$$stat_{a,t,c} = x_a + x_t + x_c + \epsilon_{a,t,c} \quad (1.9)$$

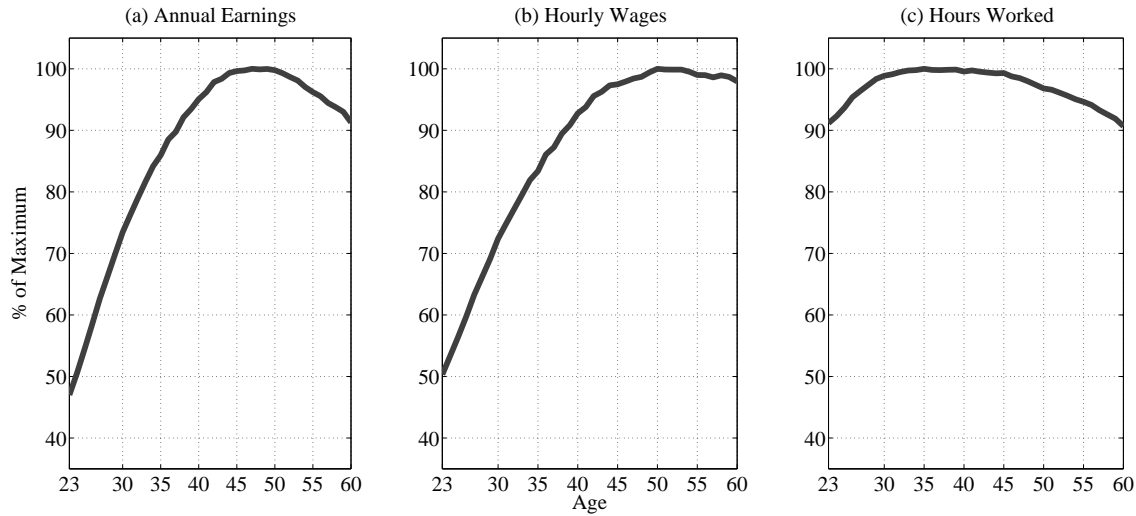
where a, t, c denote age, year, and birth cohort, respectively; where x_a, x_t, x_c are age, year, and cohort fixed effects, respectively; and where $\epsilon_{a,t,c}$ is an error term. Since $c \equiv t - a$, there is a well-known multicollinearity problem involved with estimating this model. In light of this I assume that $x_c = 0 \forall c$ (a “year-effects” view).

1.3.1 Life-cycle Profiles Targeted During Calibration of LBD and BP Models

Figures 1.1 and 1.2 depict the data moments used to calibrate the BP and LBD versions of the human capital model. Figure 1.1 plots the age effects for mean (a) annual earnings, (b) hourly wages, and (c) hours worked. Annual earnings double between 23 and 48 years of age, then decline 10% by age 60. Hourly wages double by age 50 and remain near this peak through age 60. By contrast, mean hours are relatively flat: they grow about 10% from age 23 to age 35, then gradually fall back to their original level by age 60. In absolute terms mean annual hours start at 2,011 for 23 year olds and peak at 2,200 hours.

Figure 1.1: Means for Earnings, Hourly Wages, and Hours Worked.

Note: Displays the age effects for mean earnings, hourly wages, and hours worked in the data after controlling for year effects. The curves are normalized to percentages of the maximum life-cycle value.

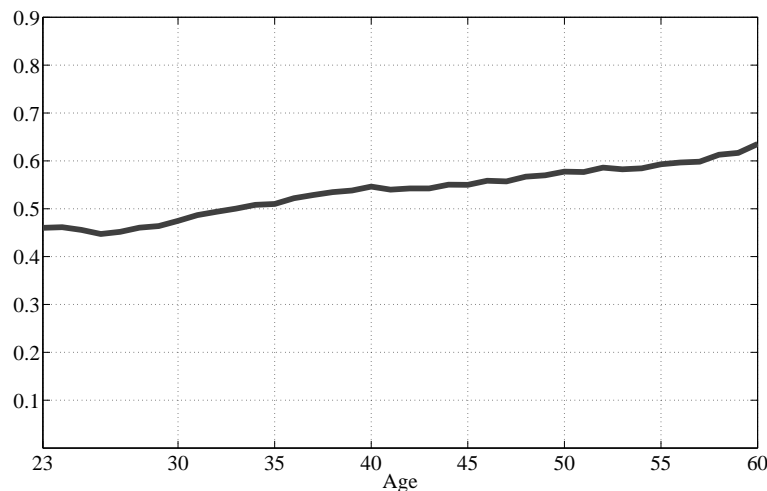


The age effects for the log variance of annual earnings are displayed in Figure 1.2. The variance in earnings among 23 year olds is approximately 45% of mean earnings, and increases to 65% of mean earnings by age 60. An important feature of this profile is that it is positively sloped at almost every age, implying earnings spread out over the life-cycle.

Another important feature is that both the initial level of the earnings variance profile, as well as the growth in the profile over the life-cycle, are primarily attributable to the variance of hourly wages. Figure B1 reveals that the initial level of the variance in hours worked is roughly a quarter of the initial level of the variance in hourly wages. Further, the variance in hours worked is roughly constant over the life-cycle, while the variance in hourly wages grows at a rate similar to that of the variance of earnings.

Figure 1.2: Life-cycle Variance of log Earnings in the Data.

Note: Displays the age effects for the variance of log earnings in the data after controlling for year effects.



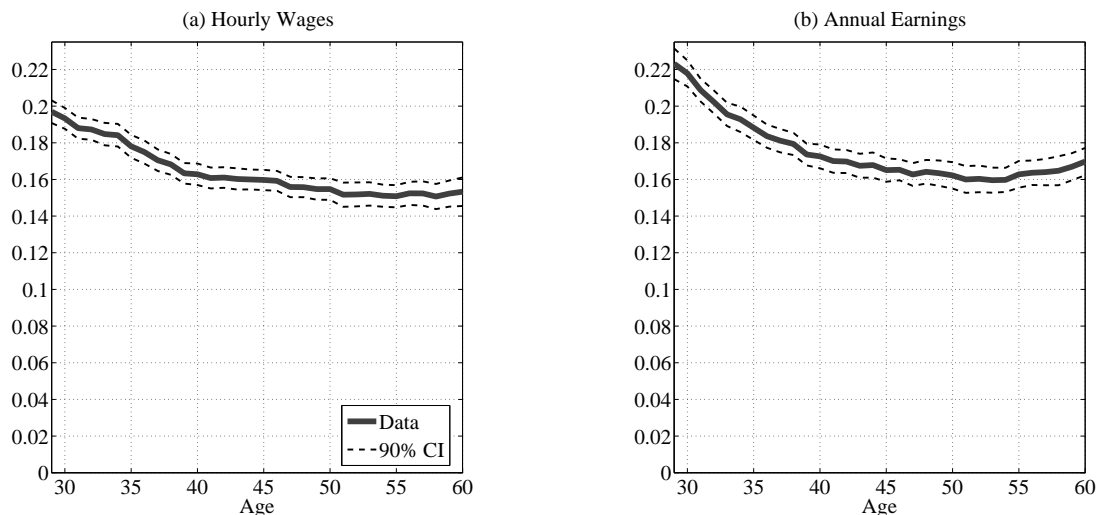
1.3.2 Untargeted Life-cycle Profiles

Figure 1.3 documents the variance in (a) wage and (b) earnings growth rates over the life-cycle. The computation of this variable merits explanation. In the calibrated human capital models a period will correspond to approximately 5 years, so I examine growth rates over 5 years in the data. To compute the wage growth statistic for age a and year t , I first group all individuals who are a years old in year t and who reported valid hourly wages at both year $t - 5$ and year t . I compute the log difference of wages between years t and $t - 5$ for each of these individuals, then find the variance in these log differences across all individuals in the group (a, t) . The statistic for earnings growth is computed analogously. Age effects for these statistics are then estimated using equation (2.18).

Figure 1.3 has two takeaways. First, there is substantial heterogeneity in the earnings growth of young US workers: the variance in earnings growth among 29

Figure 1.3: Variance of Wage and Earnings Growth in the Data.

Note: Displays the age effects for the cross-sectional variance of (a) wage and (b) earnings growth in the data after controlling for year effects. See the text for details of how this variable is computed.



year olds is .22, which implies that the standard deviation of earnings growth among these workers is 47 log points. Most of the heterogeneity in earnings growth is due to heterogeneity in hourly wage growth: the variance in wage growth among 29 year olds is .20, implying a standard deviation of 45 log points.

Second, Figure 3 reveals that the variance of 5-year wage and earnings growth rates fall by 25-30% over the lifecycle, with roughly half of this decline occurring after age 35.^{7 8} In Section 1.5 I show that the BP and LBD models laid out in Section 1.2 generate substantially different predictions for this life-cycle profile. The basic intuition is that in the BP model workers invest less as they age, which reduces the variance of earnings growth among workers with different learning abilities. In the LBD model, by contrast, workers learn via time in production, which is fairly

⁷To remove the effect of outliers, I throw out any observations with growth rates less than 1/3, and greater than 3. I follow the same procedure for the model in Section 6 when I compare the model predictions to the data. The findings in that section are robust to changes in these cutoffs.

⁸Confidence intervals are bootstrapped, resampling 1,000 times.

constant over the life-cycle (see Figure 1.1(c)). Therefore, the earnings of high and low ability workers continue to grow at different rates throughout the life-cycle, and the variance of earnings growth is relatively flat over the life-cycle.

Figures C1-C5 demonstrate that the life-cycle decline in the variance of wage and earnings growth is robust to several alternative sample restrictions and controls, including: restricting attention to full time workers; recomputing these profiles within different education groups or occupation groups; grouping individuals by years of potential experience rather than by age; restricting attention to workers who experience nonnegative wage or earnings growth; and smoothing the wages and earnings of individuals with a Moving Average process to minimize the impact of purely temporary earnings shocks. The details of these robustness checks are in Appendix C.

Summary of Key Features The key data features in Sections 1.3.1-1.3.2 are:

1. Mean hours worked are relatively flat over the life-cycle, changing by less than 10% from their initial level.
2. Mean annual earnings and hourly wages are concave, doubling from age 23-50, then declining slightly until age 60.
3. The variance of log earnings increases by about 20 log points over the life-cycle.
4. The variance of 5-year wage and earnings growth rates falls by 25-50% over the life-cycle, with half of this decline occurring after age 35.

The next sections investigate the extent to which the human capital models laid out in Section 1.2 are consistent with these features.

1.4 Setting Model Parameters

This section details how I choose parameter values for the BP and LBD models. The main idea is to set parameter values for the BP and LBD models so that each is consistent with the “standard” life-cycle moments summarized by points 1-3 in the previous section, while leaving point 4 untargeted.

First, Section 1.4.1 discusses parameters whose values are exogenously determined and identical for both models. Next, in Section 1.4.2 I jointly target the remaining parameters in the baseline BP and LBD models to the life-cycle profiles for the mean and variance of annual earnings, as well as the mean of hourly wages and hours worked. I find that the baseline BP model is more consistent with the empirical profiles than the baseline LBD model. Finally, in Section 1.4.3 I consider extensions to both the BP and LBD model in which I allow the value of model parameters to vary with the age of a worker.

When comparing hours worked and hourly wages in the model and the data, I assume reported hours worked in the data correspond to the sum of time in production and time spent investing. This is equivalent to assuming that all human capital investment takes place on the job after age 23. This assumption is standard in the BP literature (see for example [Guvenen *et al.* \(2014\)](#) or [Badel and Huggett \(2014\)](#)). My findings remain largely unchanged when I recalibrate the model assuming that up to a third of investment is not included in hours worked.

1.4.1 Parameters Set Exogenously

Solving the human capital models for a large number of individuals requires considerable computational time, so I limit the number of working periods in the model to $J = 7$. In the data I observe individuals over a 38-year working life, from ages 23-60, so each model period represents $38/J \approx 5$ years.⁹ I set $\bar{J} = 10$, so that individuals work for seven periods then live off their accumulated assets in retirement for three. Since a period corresponds to roughly 5 years this implies that retirement roughly corresponds to ages 61-75. The period real interest rate is set to $R = 1.24$ which implies an annual interest rate of 4%, and the individual discount factor is $\beta = 1/R$.

Table 1.1: Parameters Set Exogenously.

| Parameter | Description | Value |
|-----------|---|-------|
| β | Time discount factor | .81 |
| R | Period (gross) interest rate | 1.24 |
| \bar{J} | Periods in life-cycle | 10 |
| J | Periods in working life-cycle | 7 |
| γ | Curvature of utility of leisure | 2 |
| θ | Curvature of H with respect to time input | 0.7 |

I set θ , the curvature of the human capital function with respect to the time input, equal to .7, near the middle of the values used in the BP literature. Period utility

⁹I translate the annual statistics presented in Section 1.3 by partitioning the 38 years of data into J periods, then averaging the yearly statistics within each period bin. I partition ages into periods according to the following schedule: Period 1, 23-28; Period 2, 29-33; Period 3, 34-39; Period 4, 40-44; Period 5, 45-50; Period 6, 51-55; Period 7, 56-60.

over consumption and leisure takes the form

$$u_j(c, 1 - n - s) = \log(c) + \psi_j \frac{(1 - n - s)^{1-\gamma}}{1 - \gamma} . \quad (1.10)$$

I set $\gamma = 2$ to imply a mean Frisch elasticity ¹⁰ for non-leisure time of .6. The parameter values set in this section are summarized in Table 1.1.

1.4.2 Parameters Targeted to Data: No Age-Variation in Parameters

The remaining model parameters are set jointly to match the following empirical moments, presented in Section 1.3.1:

- (1-2) Period 1 and peak values of mean annual earnings,
- (3-5) Period 1, period 2, and peak values of the variance of annual earnings,
- (6) Period 4 value of mean hourly wages,
- (7) Rate of decline of mean hourly wages from periods $J - 1$ to J ,
- (8) Peak value of mean hours worked.

These parameters include:

- (a) The distribution of initial human capital and ability, $\Lambda(h_1, a)$,
- (b) The depreciation rate, δ ,
- (c) The utility of leisure parameter, ψ ,
- (d) The returns to scale of H with respect to existing human capital, ϕ .

All parameters are jointly determined, but I discuss the link between each parameter and a particular moment in the data.

Distribution of Life-cycle Earnings

The distribution of initial human capital and learning ability, $\Lambda(h_1, a)$ is closely linked to the distribution of earnings over

¹⁰See [Guvenen *et al.* \(2014\)](#). Given the utility specification above, Frisch elasticity of non-leisure time $m = n + s$ is $\varepsilon_F = \frac{1-m}{m\gamma}$. In the data, mean annual hours worked peak in period 3, at a value of 2,233. With an assumed time endowment of 5,110 hours, this implies that $\bar{m}_3^{data} \approx .44$. Since the model is calibrated to match this moment, we are left with $\gamma = \frac{1-\bar{m}_3^{data}}{.6\bar{m}_3^{data}} \approx 2$.

the life-cycle (moments 1-5). I assume $\Lambda(h_1, a)$ is log Normal, which leaves me with five distribution parameters to set: $(\mu_{h_1}, \mu_a, \sigma_{h_1}, \sigma_a, \rho_{h_1, a})$.

The mean values of initial human capital and learning ability, μ_{h_1} and μ_a , are key for mean earnings in period 1 and the peak of mean earnings, respectively. The variance of initial human capital and ability, σ_{h_1} and σ_a , are key for the variance of log earnings in period 1 and the peak of earnings variance, respectively.

The correlation of initial human capital and ability, $\rho_{h_1, a}$ is key for the variance of annual earnings in period 2. [Huggett *et al.* \(2006\)](#) show that a low correlation between initial human capital and ability will cause the spread in earnings to decline over the beginning of the life-cycle as the earnings of high- a /low- h_1 individuals catch up to the earnings of low- a /high- h_1 individuals. To be consistent with the nearly monotonic increase in the variance of earnings observed in the data the correlation parameter must be sufficiently large, so that individuals endowed with a high a also tend to have high endowments of h_1 .

Wages at the End of the Life-cycle The depreciation rate δ has a tight relationship with the decline in the mean hourly wage rate observed over the final period of the data, $\bar{w}_J^D / \bar{w}_{J-1}^D$. To understand the link between the human capital depreciation rate and the decline in hourly wages at the end of the life-cycle, recall from Section 1.2.1 that the human capital stock of a worker evolves according to

$$h_{j+1} = (1 - \delta)h_j + H(h_j, n_j, s_j; a) .$$

Therefore, the decline in a worker's hourly wage from period $J - 1$ to J , $\frac{w_J}{w_{J-1}}$, can be written

$$\frac{w_J}{w_{J-1}} = \frac{\omega h_J n_J / (n_J + s_J)}{\omega h_{J-1} n_{J-1} / (n_{J-1} + s_{J-1})} \approx \frac{h_J}{h_{J-1}} = (1 - \delta) + \frac{H(h_{J-1}, n_{J-1}, s_{J-1}, a)}{h_{J-1}}. \quad (1.11)$$

The approximate equality follows from the fact that $s_J = 0$, since the return to learning is zero in period J , and $s_{J-1} \approx 0$, which approximately holds because the return to learning is nearly zero in the period before retirement.

In the BP specification, $H \approx 0$ since $s_{J-1} \approx 0$, which means (1.11) approximately reduces to $(1 - \delta)$. Therefore the calibrated value of δ can be approximated by

$$\delta^{BP} \approx 1 - \bar{w}_J^D / \bar{w}_{J-1}^D.$$

Alternatively, in the LBD specification $H > 0$ since n_{J-1} is the relevant human capital input and is positive. Therefore, the calibrated value of δ in the LBD case will be larger than in the BP case, holding $\bar{w}_J^D / \bar{w}_{J-1}^D$ constant. Specifically, denoting $x = (h_1, a)$, the calibrated value of δ in the LBD case is given by

$$\delta^{LBD} = 1 - \bar{w}_J^D / \bar{w}_{J-1}^D + \int a h_{J-1}^{\phi-1}(x) n_{J-1}^\theta(x) d\Lambda(x) > 1 - \bar{w}_J^D / \bar{w}_{J-1}^D \approx \delta^{BP}. \quad (1.12)$$

Curvature of Mean Earnings The parameter ϕ governs the returns to scale of the human capital function with respect to existing human capital. All else equal, low values of ϕ generate large human capital growth rates early in the life-cycle and lower human capital growth rates later in the life-cycle. In the other direction, higher values of ϕ produce human capital growth rates which are more constant over the life-cycle. This means that lowering ϕ will tend to front-load wage growth towards the early part of the life-cycle. Moment (6), the value of mean hourly wages in the middle period of the working life, will therefore have a tight relationship with ϕ .

Mean Hours Worked over the Life-cycle I assume that individuals are endowed with 5110 hours per year, which translates to 14 hours per day. The utility of leisure parameter ψ is set to match the life-cycle peak of mean hours worked in the data, which is 43% of the time endowment, or approximately 2,200 hours per year.

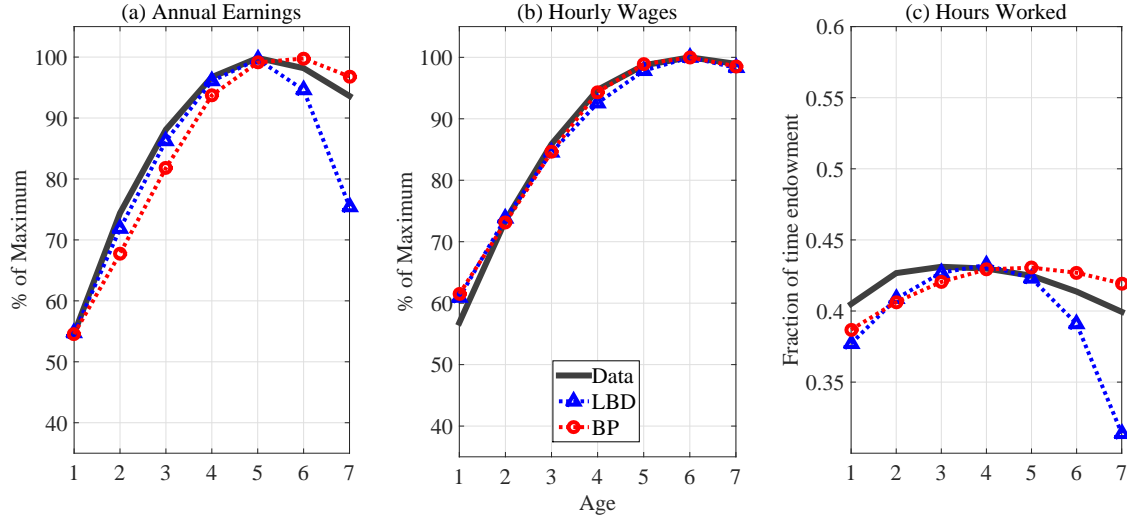
Findings: No age-variation in parameters Columns (1) and (2) of Table 1.2 summarize the calibration outcomes for the baseline BP and LBD models. Figure 1.4 plots the mean life-cycle profile of earnings, hourly wages, and hours worked for these models.

Table 1.2: Parameters Targeted to Data Moments, No Age-Variation in Parameters.

| Parameter | Description | (1) | (2) |
|----------------------------|--------------------------------------|-----------|-----------|
| | | BP: | LBD: |
| | | Basic | Basic |
| (μ_{h_1}, μ_a) | Population means for $\log(h_1, a)$ | (5.0,2.1) | (4.8,4.2) |
| (σ_{h_1}, σ_a) | Variance of $\log(h_1, a)$ | (.39,.14) | (.37,.24) |
| $\rho_{\bar{h}_1 a}$ | Correlation between (h_1, a) | .97 | .99 |
| ϕ | Curvature of $H(h, n, s)$ w.r.t. h | .47 | .2 |
| ψ | Weight on utility of leisure | 0.87 | 1.32 |
| δ | Human capital depreciation rate | .023 | .559 |

The most notable differences in parameter values are that the human capital returns to scale parameter, ϕ , is much smaller, and the depreciation rate, δ , is much larger under LBD versus BP. This is because workers in the LBD model continue to learn (via time spent producing) throughout the life-cycle: without sharply decreasing returns to scale or a high depreciation rate, LBD would counterfactually predict that

Figure 1.4: Means for the Data and the Models without Age-Varying Parameters. Note: Displays the age effects in mean earnings, hourly wages, and hours worked. The figure plots the predictions of the BP and LBD models without age-varying parameters against estimates from the data. To construct the data age profiles, I partition ages into periods according to the following schedule: Period 1, ages 23-28; Period 2, 29-33; Period 3, 34-39; Period 4, 40-44; Period 5, 45-50; Period 6, 51-55; Period 7, 56-60.



mean hourly wages continue to grow even among older workers. By contrast, workers in the BP model *choose* to stop investing as they age and the return to learning falls (see Section 1.5 for a more detailed discussion). Therefore, BP is able to produce a decline in hourly wages at the end of the life-cycle even though ϕ is closer to one and δ is close to zero.

The most notable difference in life-cycle predictions is that the LBD and BP models generate substantially different paths for mean hours worked over the second half of the life-cycle. Specifically, period J hours in the BP model are 3% below their peak, and period J hours in the LBD model are 27% below their peak, relative to a 7% decline in the data.¹¹

¹¹The large declines in hours worked predicted by the LBD model are similar in magnitude to previous predictions in the LBD literature: for example, see Imai and Keane (2004).

The first order condition in (1.8), which I partially restate below, explains why the hours decline in the BP model is small:

$$\underbrace{-\frac{\partial u_j}{\partial n_j}}_{\text{MV(Leisure)}} = \underbrace{\lambda \omega h_j}_{\text{MV(Earning)}} .$$

Hourly wages decline in the final periods of the model, which lowers the $MV(Earning)$ term on the right hand side of the above equation. In response, leisure increases in order to restore the first order condition between leisure and production. However, the decline in wages in the data, which the model has been calibrated to reproduce, is only 2%. Therefore the increase in leisure in the model, and corresponding decrease in hours worked, is small.

Alternatively, the hours decline in the basic LBD model is much larger than what is observed in the data. The first order condition in (1.7), which I restate below, shows why this is the case:

$$\underbrace{-\frac{\partial u_j}{\partial n_j}}_{\text{MV(Leisure)}} = \underbrace{\lambda \omega h_j}_{\text{MV(Earning)}} + \underbrace{\lambda \left(\frac{\partial h_{j+1}}{\partial n_j} \right) \left[\sum_{k=1}^{J-j} R^{-k} \omega n_{j+k} \left(\frac{\partial h_{j+k}}{\partial h_{j+1}} \right) \right]}_{\text{MV(Learning)}} .$$

As workers age in the LBD model, leisure increases not only in response to falling wages, but also in response to falling marginal values of learning. The marginal value of learning falls as workers approach retirement because the number of periods remaining before retirement, $J - j$, during which newly produced human capital can be used, is declining to zero. This second term is not present in the production/leisure first order condition for the BP model because individuals do not accumulate human capital through production in this model.

1.4.3 Parameters Targeted to Data: Allowing Age-Variation in Parameters

Next I show that allowing some model parameters to vary with age is sufficient to bring the profile of mean hours in both human capital models in line with the data. I first allow the utility of leisure to evolve at a constant rate over the life-cycle, $\psi_1 = \psi$ and $\psi_{j+1} = \psi_j(1 + g_\psi)$. This change affects the $MV(Leisure)$ terms in (1.7) and (1.8) for older workers. All else equal, larger values of g_ψ increase the marginal value of leisure among older workers relative to younger workers, which tends to generate larger life-cycle declines in hours worked. Put simply, assuming that older workers value leisure more will tend to increase the size of the hours decline in the run up to retirement.

Second, I impose $\psi_j = \psi \forall j$, but allow the depreciation rate to evolve at a constant rate, $\delta_1 = \delta$ and $\delta_{j+1} = \delta_j(1 + g_\delta)$. Third, I impose $\psi_j = \psi$ and $\delta_j = \delta \forall j$, but allow the learning ability of individuals to evolve at a constant rate, $a_1 = a$ and $a_{j+1} = a_j(1 + g_a)$. These latter two changes affect the $MV(Earning)$ and $MV(Learning)$ terms in (1.7) and (1.8) for older workers. All else equal, larger values of g_δ or smaller values of g_a decrease the marginal value of learning among older workers, which tends to generate smaller life-cycle declines in hours worked in the LBD model. Put simply, assuming that the value of learning for older workers is small well before retirement—either because they struggle to learn anything new or because they forget most of what they do learn—will tend to dampen the decline in hours in the run up to retirement.

In all three cases, after I add one of the parameters g_ψ , g_δ , or g_a , I introduce an additional empirical target: mean hours worked in the final working period J . The

calibration process for all other parameters remains the same as was described in Sections 1.4.1 and 1.4.2.

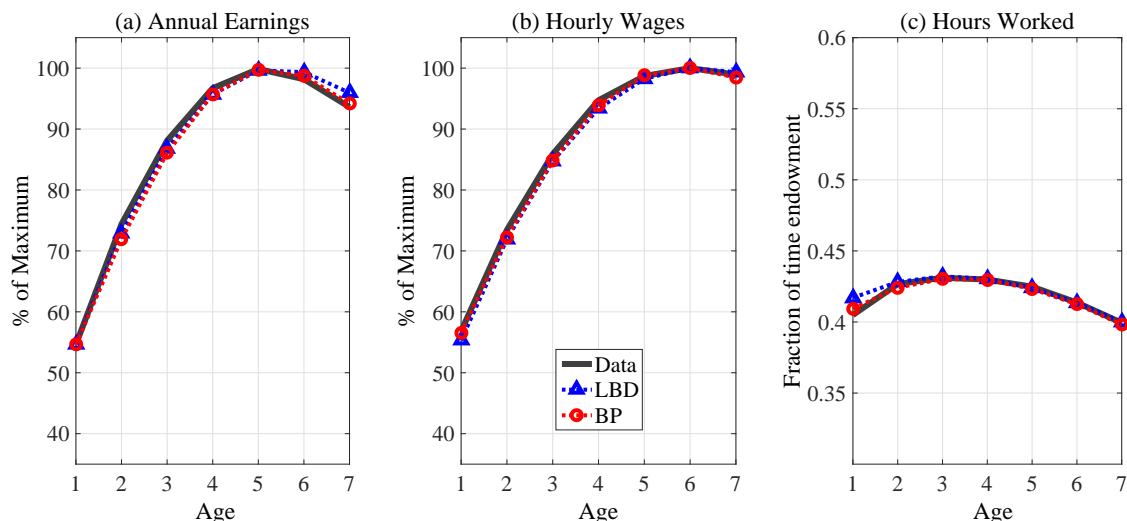
Findings: Allowing for age-variation in parameters Columns (1)-(3) in Table 1.3 summarize the calibration outcomes for the BP and LBD models with age variation in model parameters. Figure 1.5 plots the mean life-cycle profile of earnings, hourly wages, and hours worked for the models in columns (1) and (3). Figure 1.6 plots the life-cycle profile of the log variance of earnings for these models.

Table 1.3: Parameters Targeted to Data Moments, with Age-Variation in Parameters.

| Parameter | Description | (1) | (2) | (3) |
|----------------------------|--------------------------------------|-----------|------------|-----------|
| | | BP: | LBD: | LBD: |
| | | g_ψ | g_δ | g_a |
| (μ_{h_1}, μ_a) | Population means for $\log(h_1, a)$ | (5.0,2.1) | (4.8,4.0) | (4.7,2.4) |
| (σ_{h_1}, σ_a) | Variance of $\log(h_1, a)$ | (.39,.14) | (.42,.52) | (.38,.24) |
| $\rho_{\bar{h}_1 a}$ | Correlation between (h_1, a) | .94 | .94 | .68 |
| ϕ | Curvature of $H(h, n, s)$ w.r.t. h | .47 | .0 | .4 |
| ψ | Weight on utility of leisure | .79 | 1.21 | .97 |
| δ | Human capital depreciation rate | .023 | .035 | .079 |
| $(1 + g_\psi)^{J-1}$ | Lifetime growth in leisure utility | 1.22 | 1 | 1 |
| $(1 + g_\delta)^{J-2}$ | Lifetime growth in depreciation | 1 | 4.9 | 1 |
| $(1 + g_a)^{J-2}$ | Lifetime growth in ability | 1 | 1 | 0.26 |

Figure 1.5(c) shows that when I allow for age-dependence in parameters the models generate mean hours profiles which are quantitatively consistent with mean hours in the data. The BP model featuring a 22% life-cycle increase in the value of leisure

Figure 1.5: Means for the Data and the Models with Age-Varying Parameters. Note: Displays the age effects for mean earnings, hourly wages, and hours worked. The figure plots the predictions of the BP model featuring age-varying leisure preferences and the LBD model featuring age-varying learning ability against estimates from the data. To construct the data age profiles, I partition ages into periods according to the following schedule: Period 1, ages 23-28; Period 2, 29-33; Period 3, 34-39; Period 4, 40-44; Period 5, 45-50; Period 6, 51-55; Period 7, 56-60.



reproduces the 7% decline in mean hours observed in the data. In the LBD model allowing the depreciation rate of human capital to increase with age, or the learning ability of workers to decrease with age, brings the hours decline in line with the data.

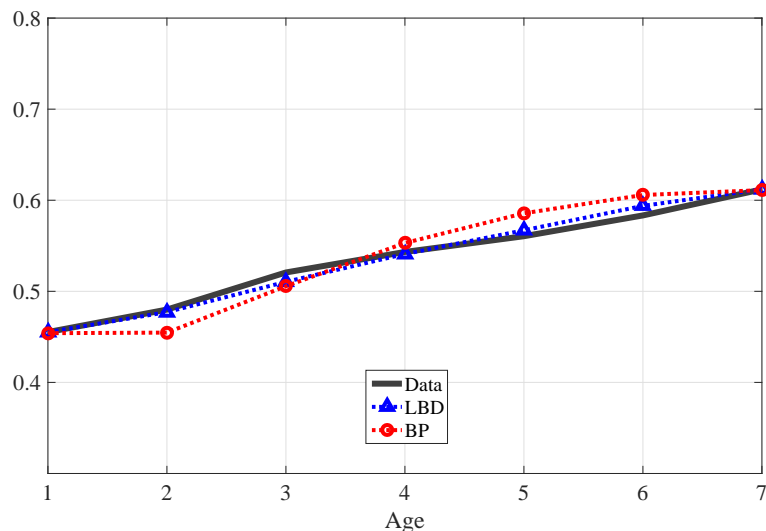
12

The primary takeaway from this exercise is that to be consistent with the distribution of earnings, wages, and hours in the data, the LBD model must feature substantial life-cycle changes in some parameter. In particular, from their early twenties

¹²At first glance the simplest way of bringing the LBD prediction in line with the data would be to make the opposite change, i.e. to decrease the value of leisure for older workers. However, increasing hours worked in this manner increases human capital accumulation late in the working life (since workers learn via production in the LBD model), which *ceteris paribus* would generate a counterfactual wage increase among older workers. The human capital depreciation rate which is required to eliminate this wage increase is in excess of 60% per year, which is an order of magnitude larger than depreciation rates typically used in the literature. Therefore I do not consider this to be a reasonable solution.

Figure 1.6: Variance of log Earnings for the Data and the Models with Age-Varying Parameters.

Note: Displays the age effects for the log variance of earnings. The figure plots the predictions of the BP model featuring age-varying leisure preferences and the LBD model featuring age-varying learning ability against estimates from the data. To construct the data age profiles, I partition ages into periods according to the following schedule: Period 1, ages 23-28; Period 2, 29-33; Period 3, 34-39; Period 4, 40-44; Period 5, 45-50; Period 6, 51-55; Period 7, 56-60.



to their late fifties individuals must experience either a 390% increase in their human capital depreciation rate or an 74% decline in their learning ability (see Table 1.3). By contrast, the BP model with no age variation in parameters matches the empirical profiles fairly well, and the BP model with a small (22%) lifetime increase in the value of leisure matches the profiles extremely well.

1.5 Earnings Growth in Ben-Porath and Learning by Doing

Section 1.4 found that the baseline BP model was more consistent with empirical profiles for earnings, wages, and hours *levels* than the baseline LBD, but that these profiles were not sufficient to differentiate between richer versions of BP and LBD. In this section, I ask whether an additional set of moments—the life-cycle declines in the

variance of earnings growth rates—is sufficient to differentiate BP and LBD. I show that the calibrated BP version of the model accounts for nearly all of the life-cycle decline in the variance of wage and earning growth rates, while the calibrated LBD models account for less than half of this decline.

Figure 1.7 plots the age profiles for the variance in earnings growth rates generated by the BP and LBD models. The dashed red (circles) line corresponds to the BP calibration in Column (3) of Table 1.3; dashed blue lines correspond to the LBD calibration for Column (4) (squares) and (5) (triangles). The data profile corresponds to Figure 1.3.¹³

I normalize the initial value of all four profiles in Figure 1.7 to zero, and focus on the change in these profiles over the life-cycle. This normalization has two benefits. First, because I normalize the level of the data profile I do not need to address measurement error (as long as the error is not age-dependent). Second, as I document in detail in Appendix C, imposing different sample restrictions on the data generate drastically different levels for the variance of earnings growth, but have little impact on the life-cycle decline in the profile.

The key takeaway from Figure 1.7 is that the BP profile accounts for roughly 90% of the life-cycle decline observed in the data, while the LBD profiles account for between 11% and 31% of the life-cycle decline.

¹³The 5-year wage growth rates in Figure 1.3 approximately correspond to one-period wage growth rates, since a period represents $38/7 \approx 5.4$ years. Therefore I compare age effects for these growth rates against 1-period growth rates produced by the model. The moments for 6-year wage growth rates are very similar to those in Figure 1.3.

Figure 1.7: Variance of 1-Period Earnings Growth Rates in the Models and the Data.

Note: Displays the age effects for the cross-sectional variance of earnings growth rates. The figure plots the predictions of the BP model featuring age-varying leisure preferences (circles), the LBD model featuring age-varying learning ability (triangles), and the LBD model featuring age-varying depreciation rates (squares) against estimates from the data. To construct the data age profiles, I partition ages into periods according to the following schedule: Period 1, ages 23-28; Period 2, 29-33; Period 3, 34-39; Period 4, 40-44; Period 5, 45-50; Period 6, 51-55; Period 7, 56-60.

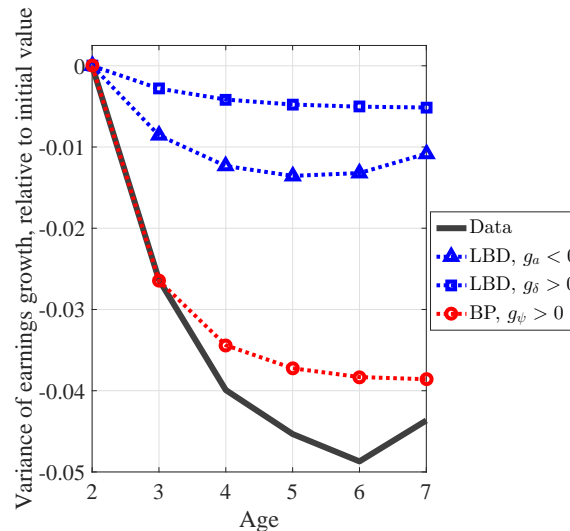
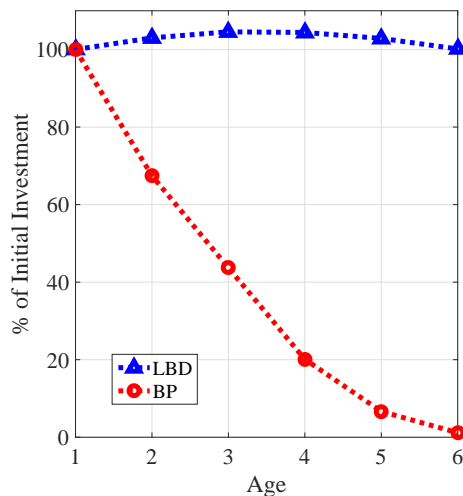


Figure 1.8 demonstrates why BP predicts a large decline in the variance of earnings growth, but LBD does not. With BP, time spent learning declines to zero as workers approach retirement. Therefore, while the earnings of young high ability workers grow much faster than the earnings of young low ability workers, this growth gap dies out over the life-cycle as workers stop investing. In the LBD model, on the other hand, time spent learning is flat over the life-cycle, as shown by Figure 1.8(a). This means that the earnings of high ability workers continue to grow *relatively* faster than the earnings of low ability workers throughout the life-cycle, even though the *average* rate of earnings growth falls as workers age.

Summary of Life-cycle Predictions

Sections 1.4-1.5 compare the life-cycle

Figure 1.8: Mean Time Spent Learning in Ben-Porath and Learning by Doing.
 Note: Plot displays age effects for mean time spent learning in the BP and LBD models corresponding to columns (1) and (3) of Table 1.3, respectively.



predictions of the BP and LBD models to data on US workers. I find that a BP model (with little or no age variation in model parameters) endogenously produces life-cycle patterns which are quantitatively consistent with the empirical distribution of earnings, wages, and hours *levels*, as well as with heterogeneity in earnings *growth rates*. I find that the LBD model is consistent with the empirical levels profiles (if I allow model parameters to vary with age), but that it is inconsistent with heterogeneity in earnings growth.

The models predict different earnings growth patterns because workers in the BP model choose to spend less time learning as they age and the return to learning falls, while workers in the LBD model choose to keep time spent learning constant with age. The intuition is that workers in the BP model can separately adjust time spent learning and earning as the returns to these activities change over the life-cycle, while workers in the LBD model cannot.

Section 1.2.3 argued that this same intuition explains why the models predict different outcomes when the relative returns to earning and learning are altered by a change in government policy. Specifically, a policy change that lowers the return to learning relative to the return to earning causes workers in a BP model to spend less time learning and more time earning; in the LBD model, by contrast, both activities will move together one-to-one. In the final section of this paper I apply these insights to counterfactual increases in the top marginal labor tax rate for the US. I find that the quantitative impact of policy crucially depends on whether a BP or LBD human capital process is assumed.

1.6 The Impact of Tax Progressivity on Human Capital Accumulation for Ben-Porath and Learning by Doing

In this section I quantitatively analyze the response of workers in the BP and LBD models to changes in the progressivity of labor income taxes. I focus on two separate tax changes. Section 1.6.1 studies an increase in marginal tax rates imposed on the highest 1% of earners. Section 1.6.2 studies a revenue-neutral change from the current (progressive) US income tax to a proportional scheme.

My analyses are complementary to work by [Guvenen *et al.* \(2014\)](#), and [Badel and Huggett \(2014\)](#). These studies use a Ben-Porath model to show that a progressive tax on earnings can generate quantitatively large reductions in human capital investment among high earners, leading to large reductions in the lifetime earnings of these individuals. The intuition is that, because earnings tend to increase over the course of a worker's life, under a progressive earnings tax workers typically expect to face a higher marginal tax rate in the future than they experience this year. Since the

return to learning depends on future marginal tax rates, while the return to earning depends on current marginal tax rates, progressive tax schemes lower the return to learning, relative to the return to earning, for most workers, which discourages human capital investment.

A question that the above authors do not address is whether their choice of a specific human capital technology is quantitatively important. I find that the learning and earning decisions of workers in the BP model respond much more to the tax changes than those in the LBD model. This suggests that the authors' choice of a BP technology is a key driver of their findings.

1.6.1 Increasing Marginal Tax Rates on the Top 1% of Earners

I conduct the tax experiments within the BP model featuring age variation in leisure preferences ($g_\psi > 0$), and the LBD model featuring age variation in learning ability ($g_a < 0$), as these are the versions of each model which are most consistent with the life-cycle earnings and hours data for US workers. I use the parameter values presented in columns (1) and (3) of Table 1.3, respectively.

The baseline government policy is modeled after the 2009 US federal labor income tax regime. The statutory average tax curve can be approximated by the logarithmic function ¹⁴ $\tau(y/\bar{y}) = .1129 + .0355\log(y/\bar{y})$. To avoid modeling the complex set of deductions which are available to households I simply assume that all households are

¹⁴Guner *et al.* (2014) provide parametric estimates of tax functions which approximate federal income tax liabilities for US households. Their estimates correspond to liabilities as a function of *total* household income. To isolate the effects of a change in the *labor income* tax I opt to use statutory federal tax rates for labor income, but make convenient use of their tax function to approximate these rates.

able to avoid one third of this tax. With this assumption, the tax schedule I use in the model closely resembles the empirical findings by [Guner *et al.* \(2014\)](#) based on administrative data about actual taxes paid by a large sample of US households.

Figure 1.9: Schedule of Earnings Taxes.

Note: The figure displays schedule of (a) average, and (b) marginal earnings tax rates as a function of earnings. Earnings are expressed as multiples of the mean earnings level in the economy. The solid line depicts the US statutory tax schedule. The short dashed line depicts the approximation of this schedule by a function which is linear in the log of mean earnings. See text for details on these tax functions.

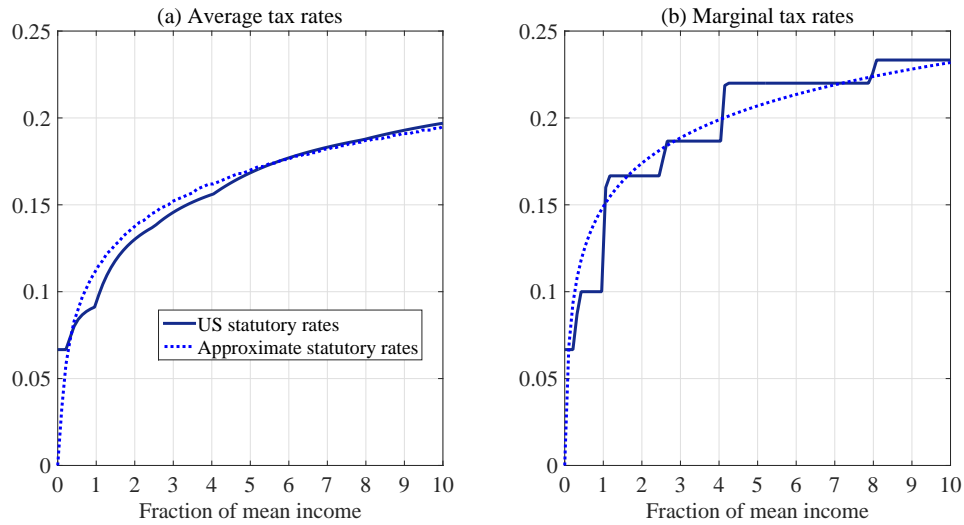


Figure 1.9 plots the resulting (a) average and (b) marginal taxes as a function of multiples of the mean labor income of households in the 2009 sample of the PSID. The parametric tax function, and the corresponding marginal tax function, are also plotted in Figure 1.9. The plot shows that workers with labor earnings equal to one tenth of mean labor earnings in the economy face a marginal tax rate of 7%; this rate rises monotonically to 23% for workers whose labor earnings are ten times mean labor earnings. I assume that in the baseline policy the government rebates all revenue back to workers lump sum. When computing the equilibrium I assume a small open economy and no population growth.

I analyze the steady state impact of increasing marginal tax rates for all earnings above the 99th earnings percentile in the baseline economy. Specifically, in both the LBD and BP models I find the level of earnings $e^{1\%}$ such that 1% of workers have earnings above this level: in the BP model $e^{1\%}$ is 4.0 times mean earnings in the economy, while in the LBD model $e^{1\%}$ is 4.3 times mean earnings in the economy. I then replace the marginal tax rate imposed on all earnings above $e^{1\%}$ with a 73% marginal rate. [Diamond and Saez \(2011\)](#) use a static model with exogenous human capital to argue that this is the marginal tax rate which maximizes the revenue the government can collect from the top 1% of earners in the US. To isolate the impact of the tax change I ignore government budget balancing considerations after the tax change.

Findings Table 1.4 summarizes the steady state change in hours worked, time spent learning, earnings, and revenue contributed by workers who earned above $e^{1\%}$ at some point in their working life in the baseline economy. The change in hours is similar between BP and LBD, 13% versus 16%. However, the change in time spent learning is drastically different between the two models. In the LBD model time spent learning (production) falls by 16%. Notice that the change in time spent learning in the LBD model is the same as the change in hours worked: this is because in the LBD model time spent learning and time spent working move together one-to-one by construction. By contrast, in the BP model workers can separately *reduce* investment (since the return to investing is now relatively lower for high earners) and *increase* time in production (since the return to earning is now relatively higher for high earners). As a result, time spent learning under BP falls nearly three times as much as under LBD: 44% versus 16%.

Table 1.4: Steady State Impact of the Tax Change on the Top 1% of Earners.

| | % Change, Time Spent Learning | % Change, Hours Worked | % Change, Earnings | % Change, Taxes Paid |
|--------------|----------------------------------|---------------------------|-----------------------|-------------------------|
| BP | -43.8% | -13.2% | -38.3% | -36.4% |
| LBD | -16.1% | -16.1% | -27.0% | -7.4% |
| Ratio | 2.7 | 0.8 | 1.4 | 4.9 |

The larger decrease in time spent learning in the BP model causes the average stock of human capital among high earners to decrease more in the BP model. This has important consequences for the earnings of these workers, and the quantity of taxes they pay. Earnings among affected workers fall by 38% in the BP model, which is 42 percent, or 11 percentage points, larger than the decline generated by the LBD model. Taxes paid by top earners decline by 36% in the BP model but only 7% in the LBD model. The reason the drop in taxes paid for the LBD model is small relative to the drop in earnings is because some workers in the LBD model continue to earn above $e^{1\%}$ after the tax change, and so are subject to the increased marginal tax rate. By contrast, in the BP model investment falls by so much after the tax change that effectively no workers pay the higher marginal rate.

In short, the long run impact of an increase in marginal tax rates on high earners crucially depends on whether a BP or LBD model is assumed.

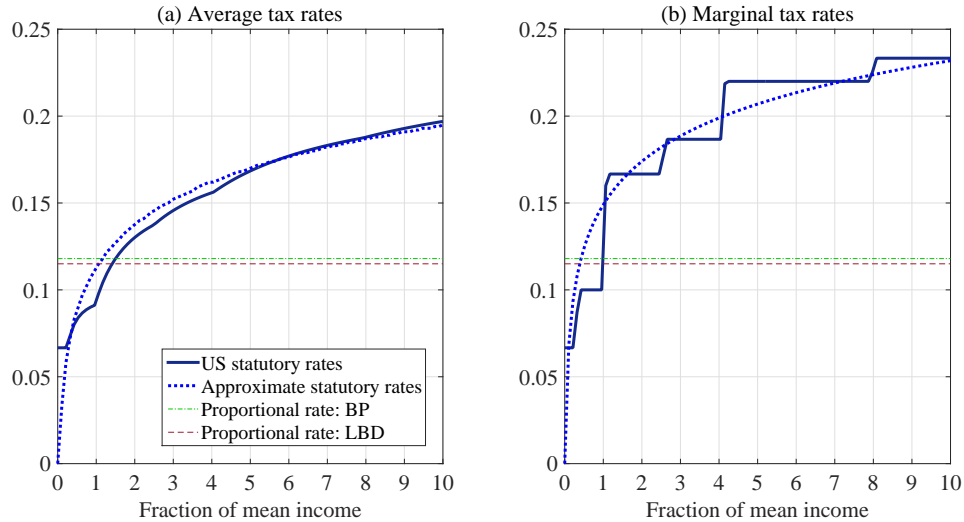
1.6.2 Change from the Current US tax scheme to a Proportional Scheme

Next I analyze the steady state impact of a more broad-based tax change, from the 2009 US federal earnings tax regime to a proportional earnings tax. In contrast

to the tax reform in the previous section, this tax change tends to *raise* the return to learning relative to earning for most workers. The reason is that after the tax change workers will face identical marginal tax rates on their future and current income, while before the tax change workers tended to face higher marginal tax rates on future income relative to current income (since income increases over the life-cycle for most workers).

Figure 1.10: Schedule of Earnings Taxes.

Note: The figure displays schedule of (a) average, and (b) marginal earnings tax rates as a function of earnings. Earnings are expressed as multiples of the mean earnings level in the economy. The solid line depicts the US statutory tax schedule. The short dashed line depicts the approximation of this schedule by a function which is linear in the log of mean earnings. The dot-dashed line depicts the proportional tax rate imposed on the BP model; the long-dashed line depicts the proportional tax rate imposed on the LBD model. See text for details on these tax functions.



Like the previous section, I conduct the tax experiment within the BP and LBD models displayed in columns (1) and (3) in Table 1.3. The proportional tax rate is set so that the revenue it generates is equal to the revenue generated by the original US tax scheme: the resulting values are an 11.8% rate in the BP model and a 11.5% rate in the LBD model. Figure 1.10 displays the initial US tax scheme, as well as

the proportional tax schemes for the BP and LBD models. To isolate the impact of the earnings tax change on worker decisions I assume that in both tax regimes the government rebates all tax revenue to households in a lump sum transfer. When computing aggregate values I assume a small open economy and no population growth.

Findings Table 1.5 summarizes the steady state change in economic aggregates induced by the tax change. The most striking effects of the tax change are with respect to the first column: switching from a modestly progressive labor income tax to a proportional labor income tax increases aggregate time spent learning (investment) by 11.7% in a BP economy. The size of this increase is six times the change in aggregate time spent learning (production) in the LBD economy. Aggregate earnings in the BP economy increase by 3.7%, which is 1 percentage point, or 37 percent, larger than the change in the LBD economy.

Table 1.5: Steady State Impact of the Tax Change on Aggregate Variables.

| | % Change, Time Spent Learning | % Change, Hours Worked | % Change, Earnings |
|--------------|--|-----------------------------------|-------------------------------|
| BP | +11.7% | +2.3% | +3.7% |
| LBD | +1.9% | +1.9% | +2.7% |
| Ratio | 6.2 | 1.2 | 1.4 |

To summarize, the aggregate effects of a revenue neutral change from a modestly progressive labor income tax to a proportional labor income tax are substantially larger in a BP environment than in a LBD environment.

When interpreting the results of the policy experiments in this section, it is im-

portant to remember that I have assumed the distribution of initial human capital and learning ability is policy invariant. Just as human capital after age 23 responded to changes in the progressivity of income taxes, it is likely that early human capital investment (before age 23) would also respond to these changes.¹⁵ The findings presented here can therefore be viewed as a conservative floor for the true impact of these policy changes.

1.7 Conclusion

This paper constructs a life-cycle human capital model which features either a BP or LBD technology. The central difference between BP and LBD is that in the BP case workers face a tradeoff between earning and learning, while in the LBD model earning and learning occur simultaneously. The two questions this paper asks are: First, to what extent are BP and LBD consistent with empirical life-cycle patterns of earnings and hours worked for US workers? Second, are the life-cycle predictions of these models informative about the impact of government policies on worker decisions?

My findings are as follows. First, both models are consistent with the life-cycle distribution of earnings and hours *levels* for male workers in the US. Second, I find that the LBD model requires substantial age variation in key parameters in order to replicate these patterns, while the BP model does not. Third, the BP model is consistent with the observed heterogeneity in earnings *growth* over the life-cycle, while the LBD model is not.

¹⁵For example, see [Krueger and Ludwig \(2013\)](#), who study the effect of taxation on education decisions.

Finally, I show that the model features which generate different life-cycle predictions across BP and LBD are the same features which generate different policy implications across the two models. Specifically, I show that workers in the two models respond differently to a change in the return to learning relative to the return to earning, and that these relative returns vary both over the life-cycle and across policy regimes. I demonstrate that the difference in policy implications can be quantitatively important by analyzing the impact of an increase in marginal tax rates imposed on the top 1% of earners in both models. I find that, among affected workers, steady state time spent learning decreases by 44% in the BP model relative to 16% in the LBD model as a result of this tax change. This leads to a decrease in earnings of 38% and a decrease in taxes paid of 36% for affected workers in the BP model, relative to decreases of 27% and 7%, respectively, in the LBD model. I also show that a revenue-neutral change from the current US labor income tax to a proportional scheme generates larger steady state increases in time spent learning and earnings in the BP model.

In sum, I demonstrate that a simple economic story in which workers respond optimally to a tradeoff between earnings today and earning potential tomorrow can explain a broad set of life-cycle facts for US workers. I also show that when this tradeoff exists, human capital investment decisions and earnings will be sensitive to changes in the returns to human capital investment. This suggests that policies which affect the return to human capital investment—such as changes to tax progressivity, education subsidies, or pension reform—will generate quantitatively large endogenous responses in key macroeconomic variables.

Chapter 2

REFORMING SOCIAL SECURITY: THE EARNINGS CAP AND ENDOGENOUS HUMAN CAPITAL

2.1 Introduction

Old age Social Security benefits in the US are funded by a 10.6% payroll tax up to a cap, currently set at \$118,500. In recent years roughly 15% of earnings were in excess of the cap, and therefore not subject to the payroll tax, with about 7% of covered US workers reporting some earnings in excess of the cap in a given year. During the 113th US Congress (01/2013 - 01/2015), 6 separate bills were introduced which sought to eliminate this cap. Additionally, the two leading contenders for the Democratic presidential nomination, Hillary Clinton and Bernie Sanders, have said they would consider raising or eliminating the earnings cap if they were to win the presidency.¹ Despite calls for removing the earnings cap, there has been very little work examining the impact of such a policy change. This paper quantitatively assesses the aggregate and distributional impact that changing the earnings cap would have on the US economy.

An increase in the earnings cap is likely to have important implications for the earnings of workers for three reasons. First, eliminating the cap will substantially lower the marginal after-tax wages of many high earners. Workers with earnings

¹See [Whitman and Shoffner \(2011\)](#) for details on earnings above the Social Security cap. The six bills which would eliminate the earnings cap, some of which were introduced to both the House of Representatives and the Senate, are: S.567 and H.R. 3118, S.308 and H.R. 649, S.500 and H.R. 1029, S.2455, H.R. 1374, H.R. 5306. See [Ehrenfreund and Gearan \(2015\)](#) for details on the positions of Clinton and Sanders.

above the current cap already face marginal tax rates between $25\% - 39.6\%$ due to the federal income tax. Subjecting the marginal earnings of these workers to the 10.6% payroll tax would increase these marginal tax rates to $35.6\% - 50.2\%$, which would lower the after-tax wages of these workers by $14\% - 18\%$. Such a large change in after-tax wages could potentially generate large reductions in hours by affected workers.

Second, recent work by [Guvenen *et al.* \(2014\)](#); [Badel and Huggett \(2014\)](#); and [Blandin \(2015\)](#) shows that policies which disproportionately raise marginal tax rates on high earners reduce the returns to human capital investment. Since eliminating the earnings cap increases marginal tax rates only for workers with sufficiently high earnings, this policy change could generate quantitatively large reductions in human capital investment, causing the lifetime wages and earnings of these workers to decline.

Finally, eliminating the earnings cap is likely to have a large impact on the earnings of high earners because the current earnings cap generates a non-convexity in the budget set of workers. Using a simplified model I theoretically demonstrate that an increase in marginal tax rates on high earners produces a larger decrease in labor (and therefore earnings) when I model this non-convexity, relative to an identical tax increase in a model without this non convexity. The intuition is that, with a non-convex budget, a worker can be nearly indifferent between two points on his budget line which are far apart from each other. Therefore, a small tax increase can cause a discrete decline in the worker's labor choice.

In order to evaluate the impact of a change in the earnings tax, then, several model features seem crucial. Workers should have heterogeneous earnings, so that

some are directly affected by the tax change while others are not. The government policies should include both the Social Security payroll tax and the federal income tax, to ensure that changes in after-tax wages are realistic. Finally, the labor supply and human capital of workers should be endogenous, since the tax change will impact the return to working and investing for high earners. I construct an equilibrium life-cycle model that includes these features. Individuals are born as 23 year old adults endowed with an initial level of human capital and a learning ability. These initial endowments are heterogeneous across workers. In each period workers decide how to divide their time between leisure, production, and human capital investment. Workers are subject to a borrowing constraint, but can save via a riskless asset.

The benchmark government policy against which reforms are measured is a pay-as-you-go pension system modeled after the existing Social Security system in the US, and an income tax modeled after income tax rates in the US. The benchmark model is calibrated to reproduce age profiles for mean annual earnings, hourly wages, and hours worked over the life-cycle, as well as the variance of earnings. I use my model to analyze the steady state impact of several reforms to the earnings cap.

To begin I consider the impact of eliminating the earnings cap with no other changes relative to the benchmark government policy; this can be thought of as using the additional revenue from Social Security taxes to extend the solvency of the Social Security trust fund. Next, I consider two reforms which eliminate the earnings cap but do not have any impact on the annual Social Security deficit: the first lowers the Social Security tax rate until the deficit is unchanged, while the second increases benefits by a lump sum amount until the deficit is unchanged. The latter reform

captures key elements of the Strengthening Social Security Act of 2013, a bill recently introduced to the US Senate and House of Representatives.

My main findings are as follows:

(1) Eliminating the earnings cap generates large reductions in aggregate output and consumption, between 2.1-3.1% across the three reforms I consider. Part of this decline is due to a lower physical capital stock in the economy, as households with high earnings save less in response to lower after-tax income levels. The decline in output is also driven by a lower supply of human capital. Individuals endowed with initial human capital and learning ability both in the top 10% of the population account for about half of the decline in the supply of human capital, and the entire decline in the supply of physical capital.

(2) The role of endogenous human capital is first order. When I do not allow the life-cycle human capital profiles of workers to adjust across policy regimes, the change in economic aggregates is roughly cut in half.

(3) Total federal tax revenues never increase by more than 1.2% in any reform. While eliminating the earnings cap increases revenues from the payroll tax by up to 12%, the decline in output lowers revenues derived from the federal income tax.

(4) Despite large declines in output and small increases in revenue, roughly 2/3 of workers experience an increase in welfare after policy changes which eliminate the earnings cap, for two reasons. First, in two of the reforms which eliminate the earnings cap the supply of human capital in the economy falls more than the supply of physical capital, so workers experience a small wage increase. Second, workers whose earnings never exceed the earnings cap do not face higher tax rates, and in two reforms receive either a tax cut (reform 2) or more generous old age benefits (reform 3). Welfare gains are typically small (below 1% in consumption equivalence (CE) terms)

because the additional revenue generated by eliminating the earnings cap is small. By contrast, among workers who are made worse off by eliminating the earnings cap, expected welfare loss is 2.1-2.4% in CE terms, and exceeds 9% for workers on the low end. It is somewhat surprising that roughly a third of workers are made worse off by eliminating the earnings cap, even though less than 10% of workers earn above the cap in a given year in the baseline model. The reason is that a larger share of workers, 27% of the population, earn in excess of the cap at some point in their lifetime in the baseline model. In particular, while few workers earn above the cap in their 20's, over 20% of workers in their 50's earn above the cap.

(5) Finally, I investigate the impact of *lowering* the earnings cap from its current level, but raising the payroll tax rate to leave the Social Security deficit unchanged. In contrast to the reforms that eliminate the earnings cap, which decrease output and increase welfare for most workers, I find that decreasing the cap to a level near mean earnings in the economy increases both output and welfare. Consumption increases by 0.5%, output by 1.3%, and federal income tax revenues by 2.8%. Over 95% of workers in the economy experience an increase in welfare, on average by 1.2% in CE terms. Workers with sufficiently high lifetime earnings benefit because they face a smaller tax burden and experience only a small reduction in old age benefits. Workers with low lifetime earnings experience no direct change in their Social Security benefits, though they do face a slightly higher payroll tax rate. Lastly, all workers benefit from lower interest rates and higher wage rates after the reform, as physical capital in the economy increases to a greater extent than human capital.

Background This paper relates to two strands of literature. The first strand is a set of recent papers that analyze the link between tax increases on high earners and changes in government revenue. [Guner *et al.* \(2015\)](#) use a life-cycle model with

heterogeneous workers and continuous labor supply decisions to show that increasing tax rates on the top 5% of US households can at most generate a 2% increase in total tax revenue. [Badel and Huggett \(2014\)](#) is most closely related to this project. The authors study the Laffer Curve for top earners in a heterogeneous agent model with endogenous human capital accumulation. They show that, relative to a model with exogenous wages, their model produces a Laffer Curve for top earners which is flatter and which peaks at a lower marginal tax rate. Both of these papers model Social Security benefits as a lump sum transfer which is common to all retirees, which implies that workers do not consider the impact that their earnings have on their retirement benefits, which is crucial for the analysis conducted in this paper. The tax changes in this paper are also structured differently and targeted to a larger group of earners relative to these other papers. Lastly, these papers do not consider the aggregate or distributional impact of different uses of revenue.

My paper is also related to three papers which examine reforms to Social Security within an endogenous human capital model. [Wallenius \(2013\)](#) asks what fraction of the gap in labor supply among older workers between the US and Europe can be explained by differences in Social Security rules. [Kindermann \(2015\)](#) argues that pay-as-you-go retirement programs subsidize human capital investment relative to programs which mandate personal retirement accounts. Fan, [Fan *et al.* \(2015\)](#) use a life-cycle human capital model with endogenous retirement to show that reducing Social Security benefits will increase labor force participation and human capital investment. These papers do not consider policies which change the payroll tax, or the payroll earnings cap.²

²[Fan *et al.* \(2015\)](#) do study the impact of the wholesale removal of the payroll tax. However, they do this while leaving Social Security benefits in place, to decompose the effect of the current Social Security system into its component effects, not in the context of a plausible policy evaluation.

The remainder of the paper is organized as follows. Section 2.2 lays out a simplified model which demonstrates how reforming the Social Security earnings tax cap and benefit rule impacts the labor and human capital investment decisions of workers. Section 2.3 describes the full model economy and the Social Security system in the benchmark model. Section 2.4 details the calibration of the benchmark model to the US economy. Section 2.5 analyzes the aggregate and distributional impact of changes to the earnings cap and the benefit rule. Section 2.6 concludes.

2.2 A Simple Model of Social Security

This section presents two simple models to illustrate the impact that the Social Security payroll tax and earnings cap have on the labor and human capital investment decisions of workers. Section 2.2.1 illustrates the role the earnings cap plays in the labor decision of a worker in a static model. Section 2.2.2 illustrates the role the earnings cap plays in the human capital investment decision of a worker in a two period model. I show that eliminating the earnings cap depresses labor supply and human capital investment weakly for workers with sufficiently high earnings. In Section 2.2.1 I also show that the impact of a tax increase on high earners is larger when the initial budget set of the worker is not convex: a special case of this is the elimination of the Social Security earnings cap.

2.2.1 *The Impact of the Earnings Cap on Labor*

The model is static. There is a single worker and a single consumption good c whose price is normalized to one. The worker decides how much time to spend producing, n . The worker is paid a wage ω for each unit of time he spends in

production. To minimize notation I impose $\omega = 1$. The worker has preferences over consumption and work represented by the utility function

$$u(c, n) = c + \frac{\psi n^{1/\gamma+1}}{1/\gamma + 1} ,$$

where $\psi < 0$ and $\gamma > 1$. The assumption that utility is linear in consumption is not necessary for the results in this section, but simplifies the analysis because it assumes away any income effects.

There is also a government whose sole task is to operate a Social Security system. The system has two components: an earnings tax and an old age benefit rule. Earnings are taxed at a marginal rate τ up to some earnings level \hat{e} , beyond which they are taxed at a marginal rate $\hat{\tau}$. Note that by setting $\hat{\tau} = 0$ I can replicate the earnings cap feature of the current US Social Security system. For simplicity I assume that old age benefits b are provided lump sum.

The worker chooses production time n and consumption c to maximize utility subject to his budget constraint:

$$\begin{aligned} \max_{c, n} \quad & c + \frac{\psi n^{1/\gamma+1}}{1/\gamma + 1} \\ \text{s.t.} \quad & c \leq b + n - \tau \min\{n, \hat{e}\} - \hat{\tau} \max\{n - \hat{e}, 0\} . \end{aligned}$$

As long as $\hat{\tau} \geq \tau$ the worker's problem is strictly concave. In this case the solution is unique and the first order condition which sets the marginal benefit of working equal to its marginal cost,

$$(1 - \tau \mathbf{1}_{n < \hat{e}}) (1 - \hat{\tau} \mathbf{1}_{n \geq \hat{e}}) = \psi n^{1/\gamma} , \tag{2.1}$$

is sufficient. When $\hat{\tau} > \tau$ it may be the case that the marginal benefit of working is strictly above the marginal cost when $n < \hat{e}$, and strictly below the marginal cost

when $n \geq \hat{e}$, since the marginal tax rate increases discretely at $n = \hat{e}$. In this case no value of n satisfies (2.1), and the worker will choose $n = \hat{e}$.

By contrast, when $\hat{\tau} < \tau$ (for example, when $\hat{\tau} = 0$), the problem is no longer concave. Now (2.1) is a necessary condition, but it is not sufficient and may be satisfied at two different values of n . This is because, in contrast to the previous case where $\hat{\tau} \geq \tau$, as earnings increase to \hat{e} the marginal tax rate *decreases* discretely at $n = \hat{e}$. If this tax decrease pushes the marginal benefit of working above the marginal cost, then there will be two values of n which satisfy (2.1). In this case the optimal value of n can be found by checking which value maximizes the objective.³

Results 1 and 2 below characterize the effect that the top marginal tax rate $\hat{\tau}$ has on the labor decision of workers.

Result 1 *Consider an increase in $\hat{\tau}$ from $\hat{\tau}^0$ to $\hat{\tau}'$. Denote the worker's optimal labor choice by n^0 and n' , respectively. If $n^0 \leq \hat{e}$, then $n' = n^0$. If $n^0 > \hat{e}$, then $n' < n^0$.*

Proof First consider the case where $n^0 \leq \hat{e}$. If n^0 provides weakly higher utility than any other $n > \hat{e}$ under $\hat{\tau}^0$, then n^0 must provide strictly higher utility than any other $n > \hat{e}$ under $\hat{\tau}'$. Further, if (2.1) is satisfied at n^0 under $\hat{\tau}^0$, it is also satisfied at n^0 under $\hat{\tau}'$. Therefore, if n^0 satisfies (2.1) under $\hat{\tau}^0$, $n' = n^0$. If no n satisfies (2.1) under $\hat{\tau}^0$, then no n satisfies (2.1) under $\hat{\tau}'$, so $n' = n^0 = \hat{e}$.

³In the case that both critical values of n yield the same value of the objective, the solution is not unique. However, for a given tax scheme $(\tau, \hat{\tau})$, there are a measure zero of parameter pairs (ψ, γ) for which the solution is not unique. Results 1 and 2 ignore these cases.

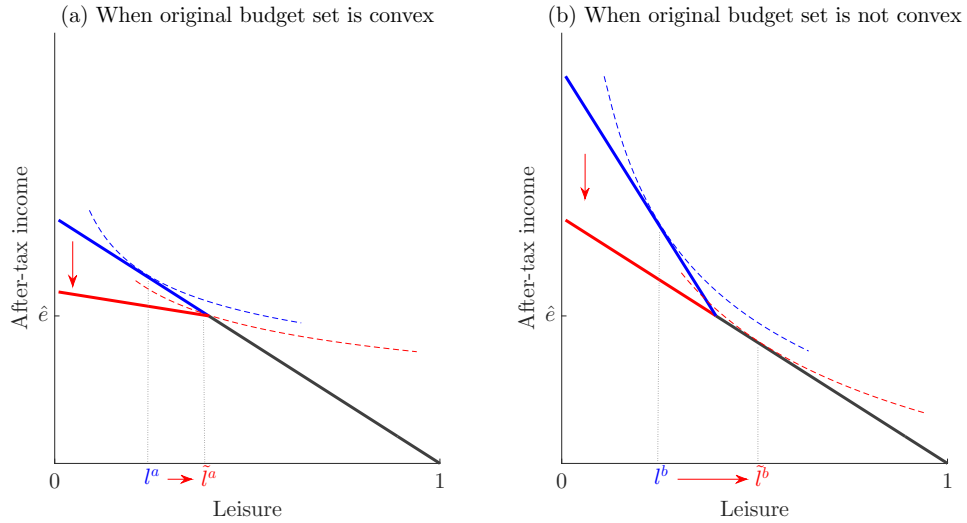
Next consider the case where $n^0 > \hat{e}$. If $n' \leq \hat{e}$, then the claim is proved, so supposed $n' > \hat{e}$. Then (2.1) is satisfied by n^0 under $\hat{\tau}^0$ and n' under $\hat{\tau}'$. But then

$$\frac{n'}{n^0} = \left(\frac{1 - \hat{\tau}'}{1 - \hat{\tau}^0} \right)^\gamma < 1,$$

so $n' < n^0$. □

Figure 2.1: Impact of a Tax Increase on High Earners.

Note: Displays the impact of a tax increase on high earners given (a) a starting budget set which is convex, and (b) a starting budget set which is not convex. See Result 2. When pretax earnings are below \hat{e} the slope of the budget line is $-(1 - \tau)$ (depicted in gray). Before the tax increase, when pretax earnings are above \hat{e} , the slope of the budget line is $-(1 - \tau^a)$ and $-(1 - \tau^b)$ in subfigures (a) and (b) respectively (depicted in blue). After the tax increase, when pretax earnings are above \hat{e} , the slope of the budget line is $-(1 - \tilde{\tau}^a)$ and $-(1 - \tilde{\tau}^b)$ in subfigures (a) and (b) respectively (depicted in red). The blue dashed indifference curves correspond to a worker's maximized utility before the tax change; the red dashed indifference curves correspond to a worker's maximized utility after the tax change.



Result 2 Consider two increases in $\hat{\tau}$: from τ^a to $\tilde{\tau}^a$, and from τ^b to $\tilde{\tau}^b$, where $\tau^b < \tau < \tau^a$. Assume that the tax increases are of the same size, i.e. $\frac{1 - \tilde{\tau}^a}{1 - \tau^a} = \frac{1 - \tilde{\tau}^b}{1 - \tau^b}$. Finally, consider two workers (ψ_1, γ) and (ψ_2, γ) , where $n_1^a = n_2^b > \hat{e}$. Then $\tilde{n}_1^b \leq \tilde{n}_2^a$,

and for some parameters the inequality is strict.

Proof First, note that $\tilde{n}_2^a \geq \hat{e}$; if not then (2.1) holds at \tilde{n}_2^a and

$$1 - \tau = \psi(\tilde{n}_2^a)^{1/\gamma} > 1 - \tilde{\tau}^a ,$$

which is a contradiction since $\tau < \tau^a < \tilde{\tau}^a$.

Second, if $\tilde{n}_1^b \geq \hat{e}$, then $\tilde{n}_1^b = \tilde{n}_2^a$. If $\tilde{n}_1^b > \hat{e}$, then (2.1) holds at both n_1^b and \tilde{n}_1^b , so

$$n_1^b \left(\frac{1 - \tilde{\tau}^b}{1 - \tau^b} \right)^\gamma = \tilde{n}_1^b > \hat{e} .$$

But then

$$n_2^a \left(\frac{1 - \tilde{\tau}^a}{1 - \tau^a} \right)^\gamma = n_1^b \left(\frac{1 - \tilde{\tau}^b}{1 - \tau^b} \right)^\gamma > \hat{e} ,$$

so

$$\tilde{n}_2^a = n_2^a \left(\frac{1 - \tilde{\tau}^a}{1 - \tau^a} \right)^\gamma = n_1^b \left(\frac{1 - \tilde{\tau}^b}{1 - \tau^b} \right)^\gamma = \tilde{n}_1^b .$$

If $\tilde{n}_1^b = \hat{e}$, then

$$n_1^b \left(\frac{1 - \tilde{\tau}^b}{1 - \tau^b} \right)^\gamma \leq \tilde{n}_1^b = \hat{e} .$$

But then If $\tilde{n}_1^b = \hat{e}$, then

$$n_2^a \left(\frac{1 - \tilde{\tau}^a}{1 - \tau^a} \right)^\gamma = n_1^b \left(\frac{1 - \tilde{\tau}^b}{1 - \tau^b} \right)^\gamma \leq \hat{e} .$$

Since $\tilde{n}_2^a \geq \hat{e}$, $\tilde{n}_2^a = \hat{e} = \tilde{n}_1^b$.

Finally, there exist values of (ψ_1, γ) s.t. $\tilde{n}_1^b < \hat{e}$, which implies $\tilde{n}_1^b < \tilde{n}_2^a$. For example, when $\tilde{\tau}^b \geq \tau$ and $\left(\frac{1-\tau}{\psi_1} \right)^\gamma < \hat{e}$ \square .

The goal of this section was to demonstrate two points. First, an increase in the tax rate on high earners $\hat{\tau}$ induces individuals to work less. An application of this

result is that eliminating the Social Security earnings cap will cause workers with high earnings to work less. Second, the impact of an increase of $\hat{\tau}$ of a given size is larger when $\hat{\tau} < \tau$ relative to when $\hat{\tau} \geq \tau$. This latter result is depicted graphically in Figure 2.1. The next section demonstrates that the first result also applies to human capital investment in a dynamic model.

2.2.2 *The Impact of the Earnings Cap on Human Capital Investment*

The model is now dynamic. There is one worker who lives for two periods, and one consumption good c whose price is normalized to one. At the beginning of the first period the worker is endowed with initial human capital h_1 and learning ability a which determines her skill in accumulating new human capital.

Workers are also endowed with a unit of time in each period. Time can be split between two activities: production, n , and human capital investment, s . The period pre-tax earnings of a worker who has human capital h and supplies n units of production are ωhn , where $\omega = 1$ is an exogenously determined wage rate set to one for simplicity. A worker who has human capital h and invests s units of time will begin the next period with

$$h' = h + ahs^\theta ,$$

where $\theta \in (0, 1)$ is a parameter which determines the elasticity of future human capital with respect to investment.

There is a government whose sole task is to operate a Social Security program, which has two components: an earnings tax and an old age benefit rule. Earnings are taxed at a marginal rate τ up to some earnings level \hat{e} , beyond which they are taxed

at a marginal rate $\hat{\tau}$. For simplicity I assume that old age benefits b are provided lump sum.

Workers have preferences over lifetime consumption (c_1, c_2) given by $u(c_1) + u(c_2)$, and can save or borrow at a risk free interest rate r . Each period workers make decisions about production n , investment s , consumption c , and savings x in order to maximize their lifetime utility, subject to a lifetime budget constraint and human capital constraints:

$$\begin{aligned}
& \max_{\{n_j, s_j, c_j, x_j\}_{j=1}^2} u(c_1) + u(c_2) \\
s.t. \quad & c_1 + x_1 = h_1 n_1 - \tau \min\{h_1 n_1, \hat{e}\} - \hat{\tau} \max\{h_1 n_1 - \hat{e}, 0\} ; \\
& c_2 = b + x_1(1 + r) + h_2 n_2 - \tau \min\{h_2 n_2, \hat{e}\} - \hat{\tau} \max\{h_2 n_2 - \hat{e}, 0\} ; \\
& h_2 = h_1 + a h_1 s_1^\theta ; \\
& n_j + s_j = 1 ; \quad x_2 = 0 .
\end{aligned}$$

A few observations greatly simplify the worker's problem. First, workers will set $s_2 = 0$ and $n_2 = 1$, since the return to human capital investment in the final period of life is zero. Second, since utility only depends on consumption and there are no borrowing/ lending constraints, the optimal time allocation decisions are simply those which maximize the present value of lifetime income. Therefore, the time allocation

decisions which solve the worker's full problem can be found by solving the simpler problem:

$$\begin{aligned} \max_{s \in [0,1]} \quad & h_1(1-s) - \tau \min\{ h_1(1-s), \hat{e} \} - \hat{\tau} \max\{ h_1(1-s) - \hat{e}, 0 \} \\ & + \left(\frac{1}{1+r} \right) [h_2 - \tau \min\{ h_2, \hat{e} \} - \hat{\tau} \max\{ h_2 - \hat{e}, 0 \}] \\ \text{s.t.} \quad & h_2 = h_1 + ah_1s^\theta . \end{aligned}$$

Result 3 shows that increasing the marginal tax rate on high earners $\hat{\tau}$ lowers the optimal human capital investment choice of the worker.

Result 3 *Consider an increase in $\hat{\tau}$ from $\hat{\tau}^0$ to $\hat{\tau}'$. Denote the worker's optimal labor choice in period 1 by s^0 and s' , and the resulting period 2 human capital by h_2^0 and h_2' , respectively. Then $s' \leq s^0$, and the inequality is strict whenever $h_1(1-s^0) < \hat{e}$ and $h_2^0 > \hat{e}$.*

Proof First consider the case where before-tax earnings in period 1 exceed \hat{e} : $h_1(1-s^0) \geq \hat{e}$. Suppose towards contradiction that $s' > s^0$, and consider the impact of increasing investment from s^0 to s' given some top tax rate $\hat{\tau}$. Doing so would decrease the worker's period 1 earnings by

$$(1-\hat{\tau}) \min\{ h_1[(1-s^0)-(1-s')] , h_1(1-s^0) - \hat{e} \} + (1-\tau) \max\{ \hat{e} - h_1(1-s'), 0 \} ,$$

and would increase the present value of the worker's period 2 earnings by

$$(1-\hat{\tau}) \left(\frac{1}{1+r} \right) (h_2' - h_2^0) .$$

In the case that $s' \geq \hat{e}$, since s^0 is optimal given $\hat{\tau} = \hat{\tau}^0$, it must be that

$$\frac{(1-\hat{\tau}^0) \left(\frac{1}{1+r} \right) (h_2' - h_2^0)}{(1-\hat{\tau}^0)[h_1(1-s^0) - h_1(1-s')]} < 1 . \quad (2.2)$$

Therefore, it must also be that (2.2) continues to hold when $\hat{\tau}^0$ is replaced by $\hat{\tau}' > \hat{\tau}^0$, which implies s' is not optimal given $\hat{\tau}'$. In the case that $s' < \hat{e}$, since s^0 is optimal given $\hat{\tau} = \hat{\tau}^0$, it must be that

$$\frac{(1 - \hat{\tau}^0) \left(\frac{1}{1+r} \right) (h'_2 - h_2^0)}{(1 - \hat{\tau}^0)[h_1(1 - s^0) - \hat{e}] + (1 - \tau)(\hat{e} - h_1(1 - s'))} < 1. \quad (2.3)$$

Since the left hand side of (2.3) would be strictly lower if $\hat{\tau}^0$ was replaced by $\hat{\tau}'$, s' cannot be optimal given $\hat{\tau}'$. Therefore, when $h_1(1 - s^0) \geq \hat{e}$, $s' \leq s^0$.

Next consider the case where $h_1(1 - s^0) < \hat{e}$. In the case that $h_2^0 \leq \hat{e}$ it must be that $s' = s^0$ since increasing investment after the tax increase can only strictly decrease the worker's lifetime earnings. Finally, consider the case where $h_1(1 - s^0) < \hat{e}$ and $h_2^0 > \hat{e}$. Then the first order condition

$$(1 - \tau) = \theta a s^{\theta-1} \left(\frac{1 - \hat{\tau}}{1 + r} \right), \quad (2.4)$$

holds at $s = s^0$ and $\hat{\tau} = \hat{\tau}^0$, which implies

$$s^0 = \left[\frac{(1 - \hat{\tau}^0)\theta a}{(1 - \tau)(1 + r)} \right]^{\frac{1}{1-\theta}}.$$

If (2.4) also holds at $s = s'$ and $\hat{\tau} = \hat{\tau}' > \hat{\tau}^0$, then $s' < s^0$. If (2.4) doesn't hold after the tax change, then it must be that $h'_2 \leq \hat{e} < h_2^0$, which implies $s' < s^0$. \square .

In this section I have presented two highly stylized models in which an increase in the Social Security taxable earnings cap depresses either labor or human capital investment for workers with sufficiently high earnings. An open question is whether the general equilibrium impact of these reforms will be quantitatively important in a richer model with more detailed tax schemes, heterogeneous agents, a life-cycle with several periods, and simultaneous labor-leisure-investment decisions for workers.

Another open question is exactly who will gain and lose from such reforms, and whose responses will be the most significant. Section 2.3 lays out a model which is capable of answering these questions. Sections 2.4 and 2.5 calibrate the benchmark equilibrium of the model to the US and analyze the impact of specific policy changes to the earnings cap and the old age benefit formula.

2.3 The Model Economy

2.3.1 *Demographics and Endowments*

I model an overlapping generations economy in discrete time. Every period a unit mass of individuals are born, each living J periods. At birth individuals are endowed with initial human capital h_1 and a learning ability a which determines their skill in accumulating new human capital. Initial endowments are heterogeneous across individuals.

Individuals are also endowed with a unit of time in each period. Time can be split between three activities: production, n , investing, s , and leisure, $1 - n - s$. The theoretical distinction between investing and production is that only time spent producing contributes to output; time spent investing is useful only to the extent that it contributes to future human capital. All individuals exogenously retire J^{SS} periods after they are born, meaning they must set leisure equal to one at ages $j = J^{SS}, \dots, J$.

2.3.2 Preferences

There is a single consumption good, c . Individuals have preferences over consumption and leisure given by the utility function

$$\sum_{j=1}^J \beta^{j-1} u(c_j, 1 - n_j - s_j).$$

Utility from leisure only depends on $1 - n - s$, and not the mix of non-leisure time spent investing or producing.

2.3.3 Human Capital Accumulation

Human capital evolves over a worker's life according to a Ben-Porath technology:

$$h_{j+1} = (1 - \delta^h)h_j + ah_j^\phi s_j^\theta.$$

A worker's human capital in period $j + 1$ is the sum of her undepreciated human capital from the beginning of period j and human capital produced in period j . δ^h is the depreciation rate of existing human capital. Newly produced human capital in period j depends on the worker's learning ability a , her existing human capital h_j , and the amount of time she spends investing in period j , s_j . The parameters $\phi < 1$ and $\theta < 1$ determine the elasticity of newly produced human capital with respect to existing human capital and investment time, respectively.

2.3.4 Technology

Output is produced by a representative firm that operates a constant returns to scale technology:

$$Y = F(K, H) = AK^\alpha H^{1-\alpha},$$

where K and H denote aggregate physical capital and human capital inputs. Total factor productivity A is assumed to be constant over time. Physical capital depreciates at rate δ^k . The firm rents physical capital and human capital from individuals in competitive markets.

2.3.5 Income Taxes

The government levies an income tax modeled after the current US federal income tax. I make use of an off-the-shelf tax function proposed and estimated by [Guner *et al.* \(2014\)](#) using data from the US Internal Revenue Service:

$$t(y/\bar{y}) = \max\{ \eta_0 + \eta_1 \log(y/\bar{y}) , 0 \} ,$$

where y is the before-tax income of the worker, \bar{y} is the average before tax income in the economy, and $t(y/\bar{y})$ is the average tax rate paid by the worker. The total income tax liability is given by

$$T(y/\bar{y}) = t(y/\bar{y})y .$$

I use the estimates they obtain from their entire US sample: $\eta_0 = .099$ and $\eta_1 = .035$.

2.3.6 Social Security

In the benchmark economy the government operates a pay-as-you-go pension system modeled after the current US Social Security system. The pension system taxes the earnings of workers up to a capped amount \hat{e} at a flat rate τ^{SS} ; earnings beyond \hat{e} are not taxed. The pension system also delivers benefits to retired workers which are a function of the workers' average lifetime earnings at the point of retirement. Specifically, workers begin period 1 with zero average earnings. A worker who began period j with average earnings \bar{e} , and who received earnings e during period j , begins

period $j + 1$ with average earnings

$$\bar{e}' = \frac{j\bar{e} + \min\{e, \hat{e}\}}{j + 1}.$$

Note that earnings beyond the earnings tax cap \hat{e} do not contribute to a worker's average earnings. Retired workers receive a benefit $b(\bar{e}_{J_{SS}})$ in each period of retirement based on their average earnings at the time they retired.

2.3.7 The Government Budget Constraint

The government raises revenue from two sources: income taxes and the Social Security payroll tax. Government expenditures are the sum of expenditures on old age Social Security benefits and government consumption, G . In the benchmark economy government consumption is set so that the government's budget is balanced in each period.

2.3.8 The Decision Problem of an Individual

Individuals in the beginning of a period are heterogeneous in five dimensions, summarized by an age j and a state vector $x = (k, h, \bar{e}, a)$, where k denotes the quantity of physical assets owned by the individual, h denotes human capital owned by the individual, \bar{e} denotes the average earnings of the individual prior to period j , and a denotes the learning ability the individual was endowed with at birth. The distribution over states x at each age j is given by the function $\Lambda_j(x)$. All individuals are born in period 1 with zero assets and zero average earnings, so $\Lambda_1(x)$ is effectively a bivariate distribution over (h_1, a) .

Given her state, an individual who is of working age decides how to split her period time endowment between production n , human capital investment s , and leisure

$1 - n - s$. If $j \geq J^{SS}$ the individual is retired and sets leisure equal to one. Individuals also decide on a quantity of physical assets to hold next period k' and how much to consume this period c .

The individual problem is solved recursively. The value function of a working age individual ($j < J^{SS}$) is given by:

$$V_j(x) = \max_{c, k', n, s} u(c, 1 - n - s) + \beta V_{j+1}(x') \quad (2.5)$$

$$s.t. \quad (1 + \tau^c)c + k' = k(1 + r) + whn - \tau^{SS} \min\{whn, \hat{e}\} ; \quad (2.6)$$

$$h' = (1 - \delta^h)h + ah^\phi s^\theta ; \quad (2.7)$$

$$\bar{e}' = \frac{j\bar{e} + \min\{whn, \hat{e}\}}{j + 1} ; \quad (2.8)$$

$$k' \geq \underline{k} ; \quad (2.9)$$

$$n, s \geq 0; \quad n + s \leq 1. \quad (2.10)$$

Equation (2.6) is the period budget constraint facing the worker, where r is the risk free interest rate that physical capital receives, w is the wage rate paid to a unit of human capital, and whn is the worker's pretax earnings. Equations (2.7) and (2.8) govern the evolution of human capital and the average earnings of the individual from period j to period $j + 1$, and equation (2.9) is a constraint on the borrowing of an individual.

Workers who are retired ($j \geq J^{SS}$) cannot work, and so face a simpler decision problem:

$$V_j(x) = \max_{c, k'} u_j(c, 1) + \beta V_{j+1}(x') \quad (2.11)$$

$$s.t. \quad (1 + \tau^c)c + k' = k(1 + r) + b(\bar{e}) ; \quad (2.12)$$

$$\bar{e}' = \bar{e} ; \quad (2.13)$$

$$k' \geq \underline{k} . \quad (2.14)$$

Equation (2.12) is the period budget constraint facing a retiree, where $b(\bar{e})$ is the benefit paid to the retiree based on her average lifetime earnings when she retired.

2.3.9 Stationary Equilibrium

This paper focuses on the long run equilibrium effects of reforms to the earnings cap and benefit rule of the US Social Security system. The benchmark economy and the reform economies are all particular parametrizations of the general model economy described in Sections 2.3.1-2.3.8. I will now define a stationary equilibrium for the general model economy.

Definition 1 *A stationary equilibrium for the closed economy is a collection of individual decisions $\{c_j(x), k'_j(x), n_j(x), s_j(x)\}$ for each age j and state x , aggregate variables $\{K, H\}$, factor prices $\{r, w\}$, government policy variables $\{\tau^{SS}, \tau^c, b, \hat{e}\}$, and measure of individuals $\Lambda(x)$ that satisfy the following conditions:*

1. *Given factor prices, individuals' decisions solve the corresponding optimization problems defined in Section 2.3.8;*

2. Factor prices are determined competitively:

$$r = F_1(K, H) - \delta^k ,$$

$$w = F_2(K, H) ;$$

3. Labor and capital markets clear:

$$H = \sum_{j=1}^J \int h n_j(x) \Lambda_j(dx) ,$$

$$K = \sum_{j=1}^J \int k \Lambda_j(dx) ;$$

4. The output market clears:

$$G + C + K = F(K, H) + K(1 - \delta^k)$$

5. The government's budget is balanced:

$$\begin{aligned} G + \sum_{j=J^{SS}}^J \int b(\bar{e}) \Lambda_j(dx) &= \tau^{SS} \sum_{j=1}^{J^{SS}-1} \int \min\{w h n_j(x), \bar{e}\} \Lambda_j(dx) \\ &+ \sum_{j=1}^J \int T(w h n_j(x)/Y) \Lambda_j(dx) , \end{aligned}$$

where $Y = F(K, H)$ is the mean income in the economy.

6. The age vector of distributions $(\Lambda_j(x))$ is stationary: For any age-state pair

$(j, x = (k, h, \bar{e}, a))$, define the law of motion for $x' = (k', h', \bar{e}', a')$ by

$$k' = k'_j(x),$$

$$h' = (1 - \delta^h)h + a h^\phi s_j(x)^\theta ,$$

$$\bar{e}' = \begin{cases} \frac{j\bar{e} + \min\{w h n_j(x), \bar{e}\}}{j+1}, & \text{if } j < J^{SS} \\ \bar{e} & \text{if } j \geq J^{SS}, \end{cases}$$

$$a' = a .$$

Then $(\Lambda_j(x))$ is stationary if, for all $(j+1, x')$,

$$\Lambda_{j+1}(x') = \int \mathbf{1}_{(x'=y')} \Lambda_j(dy) .$$

2.4 Calibration of the Benchmark Economy

This section discusses the calibration of the benchmark economy. First, I explain how I set model parameters which are not part of the Social Security policy. I then explain how I set the parameters which determine the benchmark Social Security policy.

2.4.1 Parameters Not Related to the Social Security Program

Period and life-cycle length Individuals in the model are born 23 year old adults. I assume that all individuals die at age 80. Due to large computational requirements I set the number of periods in the model $J = 12$, implying that a model period corresponds to $58/12 \approx 5$ years. I set $J^{SS} = 9$, implying that all individuals retire at age 62-67.

Technology and prices The elasticity of the aggregate production function with respect to capital is set to $\alpha = .33$ to match capital's long run income share in the US economy. The annual interest rate is set to .04, and the annual depreciation rate of physical capital is set to .0711 to generate an annual capital to output ratio of 3. This implies an annual rental rate of capital equal to $.0711 + .04 = .1111$, which translates to a period rental rate of capital equal to $R = 1.1^{58/12} = 1.6640$. The corresponding period depreciation rate of physical capital is $\delta^k = .2993$, and the

period interest rate is .3647. The wage rate is determined in equilibrium according to

$$w = F_2(K, H) = (1 - \alpha)(K/H)^{-\alpha} , \quad (2.15)$$

$$R = F_1(K, H) = \alpha(K/H)^{\alpha-1} , \quad (2.16)$$

$$\implies w = (1 - \alpha)\alpha^{\frac{\alpha}{1-\alpha}} R^{\frac{\alpha}{\alpha-1}} = .4726 . \quad (2.17)$$

The parameters governing preferences and human capital are set so that equation (2.15) holds, which closes the economy.

Preferences and human capital The remaining parameters governing individual preferences and human capital accumulation are set to replicate life-cycle moments for annual earnings, hourly wages, and hours worked in the US. For an extensive discussion of this calibration procedure, see Blandin (2015).

The distribution over initial human capital and learning abilities $\Lambda_1(h_1, a)$ is assumed to be log Normal. This leaves five parameters to calibrate: the means of (h_1, a) , (μ_h, μ_a) ; the variances of (h_1, a) , (σ_h, σ_a) ; and the correlation between h_1 and a , $\rho_{h,a}$.

The period utility function u is assumed to take the form

$$u(c, 1 - n - s) = \log(c) + \psi \frac{(1 - n - s)^{1-\gamma}}{1 - \gamma} .$$

The parameter γ determines the Frisch elasticity of non-leisure time. Specifically given a value of non-leisure time $n + s$ the Frisch elasticity is $[1 - (n + s)]/[(n + s)\gamma]$. I follow [Güvenen *et al.* \(2014\)](#) in targeting the Frisch elasticity of the average hours worked in the economy. Since average hours worked in the model are 43% of the time

endowment, I set $\gamma = 2$ to correspond with an average Frisch elasticity of non-leisure time of about 0.6.

The endowment parameters $(\mu_h, \mu_a, \sigma_h, \sigma_a, \rho_{h,a})$, the human capital depreciation rate δ^h , the leisure parameter ψ , and the discount rate β are jointly targeted to replicate seven moments in the data, and to ensure equation (2.15) holds. The data moments are: the initial and peak values of mean earnings over the life-cycle, the initial and peak values of the variance of log earnings over the life-cycle, the average life-cycle variance of log earnings, the average hours worked over the life-cycle ⁴, and the decline in hourly wages over the final period of the life-cycle.

To construct life-cycle profiles for earnings, wages, and hours I use data from the 1968-2013 family-level files of the Panel Study of Income Dynamics (PSID). I restrict my sample to observations of heads of households between 23-65 years of age who report working at least 500 hours in a given year and whose labor earnings were between \$5,000 and \$1 million in 2009 dollars. ⁵ These qualifications result in a sample of 25,245 individuals and 164,553 year-individual observations.

I restrict ages on the low end in an attempt to focus on “working age” individuals. This restriction does not imply that I ignore early-life human capital accumulation. Rather, I simply capture the result of this early accumulation in a reduced form man-

⁴Hours worked in the data are assumed to correspond to the sum of time spent producing and time spent investing. This is equivalent to assuming all investment takes place on the job, which is the standard assumption in the Ben-Porath literature.

⁵The sources of labor income used were “wages and salary”, “commission”, “overtime”, “professional practice income”, “tips”, “additional income”, and “all other labor income”. The minimum hours and earnings thresholds are designed to exclude individuals who were full time students or who experienced a long term unemployment spell. I throw out extremely high earners because these few outliers occasionally alter the measured moments.

ner, through the initial distribution of human capital and ability endowments. In fact, since these initial endowments are the only source of heterogeneity in the model, early human capital accumulation plays a crucial role in my analysis.

I now discuss how the age profiles used to set model parameters are generated. The human capital model laid out in Section 2.3 is not able to replicate every aspect of time variation contained in the data, so I use a statistical model to extract the sources of variation the model is designed to address. Following the approach of [Huggett *et al.* \(2011\)](#), I first group the data into 5-year centered age bins. I then assume that the statistics of interest are governed by the following fixed effects model:

$$stat_{a,t,c} = x_a + x_t + x_c + \epsilon_{a,t,c} \quad (2.18)$$

where x_a , x_t , x_c denote age, year, and birth cohort fixed effects, respectively, and $\epsilon_{a,t,c}$ is an error term. Since $c \equiv t - a$, there is a well-known multicollinearity problem involved with estimating this model. I assume that $x_c = 0 \forall c$ (a “year-effects” view).

2.4.2 Social Security Parameters

Two tax parameters and a benefit formula determine the Social Security policy. The parameters are the payroll tax τ^{SS} and the tax cap \hat{e} . τ^{SS} is set to 10.6% to correspond to the Social Security payroll tax rate in the US. \hat{e} is set to be consistent with the fact that 10.2% of workers in my PSID sample under the age of 65 have earnings which exceed the cap.

The old age Social Security benefit in the US is a function of a worker’s average indexed monthly earnings (AIME). To compute a worker’s AIME, the government

first gathers the reported earnings of that worker for every year of his life. The 35 years with the highest earnings are then averaged together and divided by 12 (if the worker has less than 35 years with positive reported earnings, zeros are inserted for the remaining years). Only earnings below the taxable earnings cap are used to compute this average.

An individual's monthly old age benefit is determined once and for all based off his AIME at the time the worker registers for Social Security benefits. In 2013, the benefit was 90% of the first \$791 of his AIME, 32% of the next \$4,768, and 15% beyond this point (note that benefits have a cap because only earnings below the tax cap are used to compute the AIME). The bendpoints are set in the model so that the ratio of the bendpoint to average earnings in the model is the same as the actual ratio of the bendpoint to average earnings in my 2013 PSID sample.

The values of the parameters set in this section are summarized in Table 2.1.

2.5 The Impact of Reforming Social Security

This section details the steady state outcomes of four reforms of the Social Security earnings cap. Section 2.5.1 discusses the impact of eliminating the earnings cap and using the additional revenue to reduce the Social Security funding deficit. Next I discuss two reforms which eliminate the earnings cap but do not change the Social Security funding deficit. In Section 2.5.2 the additional revenue from eliminating the cap is offset by lower revenue from lowering the payroll tax rate; in Section 2.5.3 the additional revenue from eliminating the cap is used to fund lump sum increases in old age benefits for all retirees. Finally, Section 2.5.4 lowers the earnings cap to the

Table 2.1: Model Parameters.

| | Parameter | Description | Value |
|----------------------|------------------------|---------------------------------------|------------|
| Production | α | Weight on K in production function | .33 |
| | δ^k | Annual depreciation rate of capital | .0711 |
| Preferences | β | Annual discount rate | .946 |
| | ψ | Utility of leisure | .685 |
| | γ | Frisch elasticity of non-leisure time | 2 |
| Human Capital | (μ_h, μ_a) | Mean of log (h_1, a) | (5.8, 1.6) |
| | (σ_h, σ_a) | Variance of log (h_1, a) | (.56, .35) |
| | $\rho_{h,a}$ | Correlation of log (h_1, a) | .95 |
| | δ^h | Depreciation rate of human capital | .006 |

second bendpoint level (see Section 2.4.2), and increases the payroll tax rate so that the Social Security funding deficit is unchanged.

Under the first three reforms, which eliminate the earnings cap, the maximum earnings which count towards the worker's average earnings \bar{e} are unchanged from the baseline economy. That is, although high earners now pay the payroll tax on all their earnings, earnings above the baseline earnings cap do not increase the old age benefits of those workers. This aspect of reform is in line with the Congressional bills referenced in the introduction, which call for replacement rates above the current cap which are either zero or close to zero.

To assess the role of endogenous human capital in my results, I also conduct Reforms 1-4 in a model where human capital is exogenous. To do this, I first construct a benchmark exogenous human capital model by assuming that workers face the same individual human capital profiles generated by the benchmark endogenous human

capital model, and that these human capital profiles are fixed (i.e. there is no productive role for investment in this exogenous human capital model). I then analyze the impact of Reforms 1-4 on this exogenous human capital model, *assuming that the human capital profiles of workers do not change across policy regimes.*

2.5.1 Reform 1: Eliminating the Earnings Cap and Reducing the Social Security Funding Deficit

Table 2.2: Steady State Impact of Reforms on Economic Aggregates.

| | Reform 1 | | Reform 2 | | Reform 3 | | Reform 4 | |
|---------------------------------|----------|-------|----------|-------|----------|-------|----------|-------|
| | Endog. | Exog. | Endog. | Exog. | Endog. | Exog. | Endog. | Exog. |
| Consumption | -2.9% | -1.3% | -1.8% | -0.8% | -2.3% | -1.0% | +0.5% | +0.3% |
| Output | -2.1% | -1.2% | -2.2% | -1.2% | -3.1% | -1.7% | +1.3% | +1.0% |
| Physical Capital | -1.3% | -0.9% | -1.9% | -1.2% | -3.4% | -2.4% | +2.7% | +1.8% |
| Human Capital | -2.5% | -1.2% | -2.3% | -1.0% | -3.0% | -1.4% | +0.7% | +4% |
| Hours Worked | -1.2% | -1.0% | -1.0% | -0.8% | -1.6% | -1.3% | +1.0% | +1.0% |
| Human Capital Investment | -5.1% | NA | -4.5% | NA | -5.9% | NA | -0.5% | NA |

Reform 1 sets the Social Security earnings cap $\hat{e} = \infty$, and assumes additional payroll tax revenue is spent on government consumption. This can be thought of as using the additional payroll tax revenue to improve the solvency of the Social Security system. Table 2.2 summarizes the aggregate impact of this reform. Aggregate consumption falls by 2.9%, while aggregate output falls by 2.1%. Output is lower both because savings fall by 1.3% (in response to lower after-tax income for high earners), and because the supply of human capital falls by 2.5% (see Section 2.2). The decline in the supply of human capital is due to a 1.2% decline in hours worked and to a 5.1% decline in human capital investment. Table 2.2 also makes clear that the role of endogenous human capital is first order: the change in economic aggregates

is roughly half as large when I do not allow the human capital profiles of workers to change across policy regimes.

Much of the aggregate effects of this reform can be attributed to a small group of workers with high initial human capital and high ability. Table 2.3 asks what fraction of the change in aggregate consumption, physical capital, and human capital is attributable to responses by workers who were born with initial human capital and learning ability both in the top 10% of the population. 64% of the decline in aggregate consumption, 47% of the decline in the supply of human capital, and 130% of the decline in the supply of physical capital is accounted for by these high-endowment workers.

Table 2.4 summarizes aggregate changes in government revenue and expenditures as a result of Reform 1. While revenue from payroll taxes increases by 11.8%, revenue from the federal income tax falls by 2.9%. As a result, total revenue increases by only 1.2%. This echoes recent findings by [Guner *et al.* \(2015\)](#) and [Badel and Huggett \(2014\)](#), which argue that the scope for raising additional revenue by increasing taxes on top earners is limited.

Finally, Table 2.5 analyzes the welfare implications of Reform 1. Workers benefit from higher wages and a lower interest rate, but some workers face substantially higher tax liabilities. Overall, 73% of workers are better off after the reform, while 32% are worse off. Among workers who were made worse off by the reform, average utility was 2.4% lower in CE terms. Among workers who benefited from the reform the expected welfare gain was an order of magnitude smaller, 0.2% in CE terms. The expected welfare increase is small for two reasons. First, the reform only increases

wages by 0.1% and lowers the interest rate by about one basis point. Second, the additional revenue from eliminating the earnings cap is spent on government consumption, which does not enter into the utility function of workers. Reforms 2 and 3, by contrast, use the additional revenue from eliminating the earnings cap to lower tax rates and raise old age benefits, respectively.

2.5.2 Reform 2: Eliminating the Earnings Cap and Reducing the Payroll Tax Rate

Table 2.3: Share of Aggregate Changes Attributable to High-Endowment Workers.

| | Reform 1 | Reform 2 | Reform 3 | Reform 4 |
|-------------------------|-----------------|-----------------|-----------------|-----------------|
| Consumption | 64% | 92% | 79% | 33% |
| Physical Capital | 130% | 180% | 83% | 86% |
| Human Capital | 47% | 49% | 38% | 3% |

Reform 2 sets the Social Security earnings cap $\hat{e} = \infty$, and lowers the payroll tax rate τ^{SS} so that the Social Security funding deficit (total revenue collected from payroll taxes minus total expenditures on old age benefits) is unchanged. The deficit-balancing payroll tax rate is 9.4%, down from 10.6% in the baseline economy.

Table 2.2 shows that the effects on aggregate output (-2.2%), physical capital (-1.9%), and human capital (-2.3%) are similar between Reforms 1 and 2. This is because the change in marginal tax rates for high earners is roughly the same across both reforms, and these individuals drive the changes in economic aggregates (see Table 2.3).

However, consumption only falls by 1.8% under Reform 2, relative to 2.9% under

Reform 1, because workers with earnings below the original earnings cap receive a tax decrease. This substantially improves the welfare of workers relative to Reform 1. Specifically, 78% of workers benefit from Reform 2, with an average utility increase of 1.6% in CE terms among these workers. Alternatively, 22% of workers are worse off under the reform, with an average loss of 2.1% among these workers.

While welfare is higher under Reform 2 than Reform 1, federal revenue is substantially lower because the payroll tax rate is reduced. Total federal revenues are 2.0% lower under Reform 2 relative to the baseline model, and 3.1% lower under Reform 2 relative to Reform 1.

2.5.3 Reform 3: Eliminating the Earnings Cap and Increasing Old Age Benefits

Table 2.4: Steady State Impact of Reforms on Government Revenues and Expenditures.

| | Reform 1 | Reform 2 | Reform 3 | Reform 4 |
|-------------------------------------|-----------------|-----------------|-----------------|-----------------|
| Payroll tax revenue | +11.8% | -0.5% | +11.0% | -12.7% |
| Income tax revenue | -2.9% | -2.5% | -4.5% | +2.8% |
| Total tax revenue | +1.2% | -2.0% | -0.2% | -1.5% |
| Old age benefit expenditures | -0.1% | -0.3% | +6.7% | -7.3% |

Reform 3 sets the Social Security earnings cap $\hat{e} = \infty$, and increases old age benefits for all retirees lump sum, so that the Social Security funding deficit (total revenue collected from payroll taxes minus total expenditures on old age benefits) is unchanged. The deficit-balancing lump sum benefit increase is 5.3% of the maximum possible Social Security benefit.

As Table 2.2 shows, Reform 3 generates the largest reductions in aggregate output, physical capital, and human capital among all the reforms I consider, between 3.0 – 3.4%. The reason is that relative to Reform 1 low earners work less, in response to a lump sum increase in their old age benefits; and relative to Reform 2 high earners experience an even larger increase in marginal tax rates, since the payroll tax rate is unchanged.

Reform 3 also has less positive impacts on welfare relative to Reform 2. While 63% of workers benefit from the reform, workers who benefit experience a much smaller average utility increase of 0.4% in CE terms. Workers who are worse off experience an average utility decrease of 2.3%. Fewer workers benefit, and those who do tend to benefit by less, relative to Reform 2 for two reasons. First, physical capital falls by more than human capital, so workers face lower wages and higher interest rates, both of which lower the present value of their lifetime earnings. Second, since the supply of human capital falls by more under Reform 3 than under Reform 2, the revenue base for payroll taxes falls by more, which limits the size of the lump sum benefit increase that leaves the Social Security funding deficit unchanged.

2.5.4 Reform 4: Reducing the Earnings Cap

Table 2.5: Steady State Impact of Reforms on Welfare of Newborn Workers.

| | Reform 1 | Reform 2 | Reform 3 | Reform 4 |
|--|-----------------|-----------------|-----------------|-----------------|
| Share of workers benefiting from reform | .73 | .78 | .63 | .98 |
| Welfare gain (%), conditional on benefiting | +0.1% | +1.6% | +0.4% | +1.2% |
| Welfare loss (%), conditional on losing | -2.4% | -2.1% | -2.3% | -0.1% |
| Average welfare change (%) | -0.7% | +0.9% | -0.6% | +1.2% |

In contrast to Reforms 1-3, which raised the earnings cap to infinity, Reform 4

lowers the earnings cap to the second bendpoint of the old age benefit formula. I also raise the payroll tax rate from 10.6% to 11.4% so that the Social Security deficit is unchanged. The new earnings cap is 59% of the original earnings cap. While this level is somewhat arbitrary, it is interesting for a few reasons. First, when the average indexed monthly earnings (AIME) of a worker cross the second bendpoint threshold, the replacement rates of old age benefits drops from 32% to 15%—that is, for workers who retire with an AIME beyond the second bendpoint, the marginal increase in benefits from a marginal increase in earnings is close to zero. Additionally, several OECD countries have pension systems funded by a proportional tax up to a cap (see [OECD \(2013\)](#)), and the lowest ratio of this cap to average earnings is Switzerland at 96% of average earnings. Since the second bendpoint is set to 90% of the average earnings in the baseline economy, the exercise can be thought of as replacing the earnings cap-to-average-earnings ratio in the US with a ratio close to the lowest ratio in the OECD.

Table 2.2 shows that aggregate output increases by 1.3%, aggregate physical capital by 2.7%, and aggregate human capital by 0.7%. The large increase in physical capital increases wages by 0.5% and lowers the interest rate by 10 basis points. The increase in output raises revenues from the federal income tax by 2.8%. Different from Reforms 1-3, the change in economic aggregates are not primarily driven by workers with initial endowments of human capital and learning ability in the top 10% of the population: only 33% of the increase in aggregate consumption, and 3% of the increase in aggregate earnings, are attributable to these high-endowment workers.

Most interestingly, after Reform 4 welfare increases for 95% of workers, and across all workers average welfare increases by 1.2% in CE terms. Workers experience welfare improvements for several reasons. First, many workers experience a decrease in

marginal and average tax rates: in a given period the number of workers with earnings above the earnings cap increases from 10% in the baseline model to 29% after the reform, and the fraction of workers who earn above the earnings cap at some point in their life increases from 27% in the baseline model to 56% after the reform. Second, higher wages and lower interest rates increase welfare for all workers. Third, the increase in output means that the required increase in the payroll tax rate to leave the Social Security funding deficit unchanged is small.

One note of caution is important for interpreting these results. The model is deterministic, and therefore there is no way for Social Security to provide welfare-improving insurance against random shocks to earnings or wages which might occur over the life-cycle of workers. While it would be interesting to extend this model to a stochastic setting, it is unlikely that doing so would overturn any of my key results. [Huggett *et al.* \(2011\)](#) use a Ben-Porath model to argue that 2/3 of the variance in lifetime incomes are due to initial human capital and ability differences at age 23, which is when the workers in my model are born. Additionally, workers can partially insure against income shocks by saving, which limits the insurance value of Social Security. Finally, another source of uncertainty which this model abstracts from is uncertainty over lifespans. However, [Hong and Ríos-Rull \(2007\)](#) argue that welfare-enhancing aspect of Social Security as an annuity is small relative to distortions to the intertemporal savings margin of workers, particularly in economies where life-insurance markets exist.

2.6 Conclusion

Old age Social Security benefits in the US are funded by a 10.6% payroll tax up to a cap, currently set at \$118,500. Despite calls from politicians and policy circles to eliminate the cap on taxable earnings, there has been little work examining the likely outcomes of such a policy change. I use a life-cycle model with heterogeneous individuals to investigate the aggregate and distributional steady state impacts of several policy changes the earnings cap. Three crucial features of my model are a federal income tax and social security program modeled after existing US policies; a labor-leisure decision by workers; and endogenous human capital profiles which depend on investment decisions by workers.

I find: (1) Eliminating the earnings cap generates large reductions in aggregate output and consumption, between 2.1-3.1%. (2) The role of endogenous human capital is first order: when I do not allow the life-cycle human capital profiles of workers to adjust across policy regimes, the change in economic aggregates is roughly cut in half. (3) While eliminating the earnings cap increases revenues from the payroll tax by 12%, the decline in output lowers other federal tax revenues, so that total federal tax revenues never increase by more than 1.2%. (4) Eliminating the earnings cap produces modest welfare gains for about 2/3 of workers, while about 1/3 of workers experience welfare losses, which are typically large. (5) Reducing the earnings cap to a level near the mean of earnings increases aggregate output by 1.3%, and also increases welfare for the vast majority of workers.

There are many aspects of reality that the current project abstracts from, which I leave to future work. There is no welfare-enhancing insurance role for Social Security

since the model is deterministic; introducing uncertainty in wages, human capital stocks, or lifespans could alter the impact on welfare of social security reforms. I also assume that demographics do not vary over time. Allowing population growth rates and life-spans to change over time could provide important insights into the sustainability of Social Security and the revenue impact of government reforms.

Chapter 3

EQUILIBRIUM WITH MUTUAL ORGANIZATIONS IN ECONOMIES WITH ADVERSE SELECTION

3.1 Introduction

In this paper, social insurance is defined to be the centralized provision of some form of insurance. A question is, why have social insurance? Adverse selection is often used as justification for social insurance, because many claim that adverse selection can lead to the inefficient provision of insurance by decentralized mechanisms. We find that if mutual organizations are feasible and permitted, there is no failure of decentralized arrangements in providing insurance or, for that matter, in dealing with the [Akerlof \(1970\)](#) used-car “lemons” environment or with the [Spence \(1973\)](#) job market signaling environment. If mutual insurance organizations are permitted, Adam Smith’s invisible hand works. There are no market failures due to adverse selection in providing insurance.

We modify the definition of equilibrium used by [Rothschild and Stiglitz \(1976\)](#) in essentially one respect. We permit the agents who select insurance arrangements to select mutual arrangements. A mutual arrangement is a set of contracts which specify payments contingent on both individual experiences and the aggregate experience of all contract holders. In addition, the set of contracts must be such that the organization profits are not negative. As in [Boyd *et al.* \(1988\)](#), there are mutual organizations, and a core-related equilibrium concept is developed and used. But unlike this earlier approach, the new definition of equilibrium is simpler and more general. Existence,

uniqueness (in the types' utility outcomes), and optimality are established for an arbitrary finite number of types of individuals. The earlier Boyd-Prescott equilibrium was defined only for two types. Attempts by us and others to generalize to more than two types have failed.

With the Rothschild and Stiglitz equilibrium definition, at the first stage nonmutual insurance contracts are selected. At the second stage, individuals choose their optimal contract from the set of offered contracts, or choose not to insure. In our definition, at the first stage mutual insurance contracts are selected. At the second stage, individuals pick their mutual arrangement optimally given the choices of others. Thus, the second stage is a Nash equilibrium.

We say that a mutual arrangement is blocked if there exists an alternative mutual arrangement for which a subset of types can make themselves better off using only their own resources. An equilibrium is defined to be an unblocked mutual arrangement. With our setup, if and only if an equilibrium mutual arrangement is chosen at stage 1, no agent can offer an alternative mutual arrangement that is profitable, given what will happen in the second stage.

We abstract from a number of features of reality that are important in the provision of insurance. In our examples, no costs are associated with operating an insurance company, whereas in fact operating an insurance company entails large costs. Evidence that these costs are large is that the gross output share of the U.S. insurance carriers and related activities sector reported in the U.S. National Accounts was 5 percent of GNP in 2012. Associated with the operation of an insurance business are record-keeping costs, monitoring costs needed to mitigate the moral hazard problems

associated with insurance contracts, and asset-managing costs as premiums are received prior to claims being paid. If there were a market failure, when evaluating social insurance, these costs would have to be carefully quantified and incorporated into the analysis in addition to the costs that are specific to social insurance. But there is no need to quantify all these costs because the decentralized outcome is Pareto efficient.

3.2 The Relationship to Other Equilibrium Concepts

We use [Debreu \(1954\)](#) definition of an economy. In so doing, we follow [Cornet](#), who has been a leading contributor to the development of the theory of value (see, e.g., [Cornet \(1988\)](#)). Debreu's definition of an economy requires the specification of a commodity space. The commodity space is a linear topological space. There also is a set of individuals with preference orderings on subsets of the commodity space and a set of technologies, which are subsets of the commodity space. Consequently, feasibility and Pareto optimality are well defined. In addition, the economic statistics of the national income accounts can be calculated and welfare analysis carried out.

An assumption is that the preference orderings of individuals can be characterized by the expected value of utility functions, which are continuous real-valued functions on the underlying consumption set. As in [Prescott and Townsend \(1984\)](#), the elements of the commodity space are signed measures on the Borel sigma-algebra of the underlying consumption set. The consumption sets are sets of probability measures in this space.

Equilibrium is not a Debreu valuation equilibrium as it is in the [Prescott and](#)

Townsend (1984) private information economies, where households and those who pick the commodity vector in the technology sets take prices as given. Instead, as part of the definition of equilibrium, the blocking concept in Edgeworth (1881) theory of the core is used. But groups do not block. Rather, an initial mutual organization is blocked if an active agent writes a charter for a new mutual organization which makes a subset of individuals better off while preserving resource and incentive feasibility. This is similar to competition among insurance firms in Rothschild and Stiglitz (1976), except that the choice is a mutual arrangement and not an insurance contract. Using our equilibrium definition, if mutual organizations are prohibited and nonmutual companies permitted, all the findings of Rothschild and Stiglitz (1976) hold. In their classic environment, for example, no transfers occur between individual types in equilibrium, and any equilibrium allocation must be the minimum separating allocation.

Given expected utility maximization, the utility function of each individual type is linear, and the set of incentive-compatible allocations is a convex affine cone. The intersection of the resource constraint set and the incentive-compatible allocation set is a convex subset with dimension one less than the number of types.

Each person either selects the mutual organization that is best given the choices of other people or chooses not to insure. Thus, the second stage is a Nash equilibrium. The economy is large with an atomistic measure of each of the finite number of types of individuals. Type is private information.

3.3 The Formal Analysis

The economies are large with a finite number of individual types $i \in I$. The measures of type i are λ_i . The underlying consumption set is C , a closed and bounded subset of \mathbf{R}^n . The commodity space is the space L of signed measures on the Borel sigma-algebra of C .

Each individual's consumption possibilities set is $X \subset L$, the set of probability measures on the Borel sigma-algebra of C . Preferences of a type i are ordered by the expected value of a continuous function v_i :

$$u_i(x) = \int v_i(c)x(dc).$$

An allocation is an n -tuple $\{x^i\}_{i \in I}$. The Incentive Compatibility constraints (*IC*) are

$$u_i(x^i) \geq u_i(x^j) \quad \forall i, j \in I. \quad (3.1)$$

Type i must weakly prefer x^i to x^j .

The single Resource Constraint (*RC*) has the form

$$\sum_{i=1}^I \lambda_i (r_i(x^i) - \pi_i) \leq 0, \quad (3.2)$$

where the functions $r_i : X \rightarrow \mathbb{R}$ are linear and the π_i are parameters. In the [Rothschild and Stiglitz \(1976\)](#) insurance case the resource constraint is the condition that a mutual insurance organization cannot distribute more in claims than it takes in as premiums. In the [Spence \(1973\)](#) signaling environment, it is the condition that wage payments to members of a mutual organization cannot exceed the value of production of that organization.

An allocation is feasible if it satisfies the incentive constraints, (3.1), and the resource constraint, (3.2). An allocation x is a Pareto optimum if for no other feasible allocation y , $u_i(y^i) \geq u_i(x^i)$ for all i with strict inequality for at least one i .

3.3.1 Working in the Utility Space

We transform the problem into the utility space as the proofs of existence and optimality of unblocked mutual arrangements are carried out in this space.

Associated with each allocation x is a utility vector $u(x) = \{u_i(x^i)\}_{i \in I}$. The set of utility vectors that satisfy the incentive constraints is denoted by IC . This set is a convex affine cone. The set of allocations that yield a given point in IC is convex. The mapping from utility vectors in the IC set to allocations that we use is denoted by $x(u)$. If the inverse image of $u(x)$ is not a point, then the mean allocation of all allocations that yield utility vector is the value of the function $x(u)$. This mean allocation is in the set of allocations that yield u , because the inverse image of the function $u(x)$ is a convex set.

The aggregate net transfer made by type i is

$$t_i(u) = \lambda_i(r_i(x^i(u) - \pi_i). \quad (3.3)$$

The resource constraint is

$$\sum_{i=1}^I t_i(u) = 0. \quad (3.4)$$

The set of utility vectors that satisfy this constraint is a hyper-plane with dimension one less than the cardinality of I .

Definition 2 *A feasible utility vector u is blocked by feasible u' if, for the set of types that do better under u' , denoted by B ,*

$$\sum_{i \in B} t_i(u') \geq 0.$$

3.3.2 Some Results

Result 4 *The Pareto optimum utility set, denoted P , is a closed, convex subset of the hyper-plane that satisfies (3.4).*

Result 5 *Any utility vector u which is not Pareto optimal is blocked by the set of types I and some Pareto superior feasible utility vector.*

Result 6 *Any unblocked utility vector u is a Pareto optimum.*

Result 7 *Utility vector $u \in P$ is blocked if and only if there is a $u' \in P$ that blocks it.*

Therefore, the problem is to establish the existence of an unblocked utility vector $u \in P$.

Proposition 1 *An equilibrium exists and is unique in the utility space.*

Proof Denote the Pareto utility set by P . Pick a direction in P . Moving in that direction makes one set of types better off and the other set worse off. Thus a direction partitions the types into two sets. Moving in the given direction makes one set of types better off and the complement set of types worse off.

A point u in P and a direction define a line. The intersection of a line and P is a line segment. Given a line segment in a subset in P , that subset can be partitioned into line segments parallel to this line segment. Along any of these partitions, utilities

and transfers vary linearly. A necessary condition for an unblocked utility point to be on the line is that transfers by the benefiting group for the given direction minimizes the absolute value of the transfers by that group in that partition. The set of all utility points that satisfy this property for some partition is a convex set with dimensionality less than that of the subset. All unblocked utility points lie in that set.

Given that the dimensionality of P is finite, beginning with subset P by induction a set of dimension zero can be found that contains the unblocked utility points, u^* . \square

Note that the allocation that yields u^* may not be unique. If not unique, we select the allocation that is the average of the set that yields this utility outcome. Given the linearity of the mapping from feasible arrangements to utilities, the inverse image of any point in the utility possibility set is a convex set. Thus, this average is a point in the set.

3.4 Application to the Adverse Selection Insurance Environment

The environment has measures λ_i of people of type $i \in \{1, 2, \dots, I\} = I$. The same symbol is used for both the cardinality of the set and the set itself. An individual's type is private information. A person has a random endowment subsequent to contracting $e_j \in \{1, 2, \dots, J\} = J$. The probability of a type i having endowment e_j is π_{ij} .

The underlying consumption space is $C = \{x \in \mathbb{R}_+^J : x_j \leq \bar{c} \forall j\}$. The consumption ceiling \bar{c} is sufficiently large that increasing it does not increase the set of feasible allocations. This is possible because the resource constraint, necessary for feasibility,

is bounded. The commodity point is

$$x = \{x_j(dc)\}_{j \in J}.$$

The consumption possibility set X is a vector of J probability measures, and the utility functions are

$$u_i(x) = \sum_{j \in J} \pi_{ij} \int v(c) x_j(dc), \quad \forall i \in I.$$

The function $v : C \rightarrow \mathbb{R}$ is continuous, strictly increasing, and concave.

3.4.1 The Rothschild and Stiglitz Two-Type Adverse-Selection Environment

There are two types and two possible positive endowments: $0 < e_b < e_g$. The subscripts denote bad, b , and good, g . The two individual types are low-risk, L , and high-risk, H . Thus,

$$\pi_{Lb} < \pi_{Hb}.$$

We normalize the population size to one so that

$$\lambda_H + \lambda_L = 1.$$

The average endowment for type i is

$$\bar{e}_i = \pi_{ib}e_b + \pi_{ig}e_g \quad \forall i \in I.$$

Depending on the parameter values, the equilibrium will fall into one of two categories: equilibria with no transfers between types, or equilibria with positive transfers from type L to type H individuals.

The unblocked utility vector is the one that maximizes type L utility on the utility possibility frontier. If λ_L is sufficiently large, the set of Pareto optima is a point, with

everyone consuming the average population endowment $\bar{e}(\lambda_L)$. This average is an increasing function in λ_L .

For the unblocked contract, the probability measures for an L type conditional on that individual endowment realization place all probability on a single point: c_{Lb} for e_b^L and c_{Lg} for e_g^L . This follows from the assumed properties of the utility function v . We denote the consumption of a high-risk individual, which is the same for both possible endowments, by c_H . Thus, each Pareto optimum is characterized by a triplet of real numbers (c_b, c_g, c_H) . The utility functions in terms of these three variables are denoted by $U_H(c_H) = v(c_H)$ and $U_L(c_b, c_g) = \pi_{Lb}v(c_b) + \pi_{Lg}v(c_g)$.

To find the Pareto set, we first find the utility-maximizing contract for type L given an average transfer from type L individuals to type H individuals, T . These transfers are constrained to the interval

$$T \in [0, (\bar{e}_L - \bar{e}_H)\lambda_H].$$

The upper end of the interval results in the average consumption of the two types being equal. The transfer for which the utility of type L is maximal is the equilibrium T^* .

We proceed in two steps. First, we consider the problem of maximizing the utility of a type L given average transfer T . The value of the solution is denoted by $\hat{u}_L(T)$. This program is

$$\hat{u}_L(T) = \max_{c_H, c_b, c_g} U_L(c_b, c_g)$$

subject to the following four constraints:

1. Type H consumption is feasible given T :

$$c_H \leq \bar{e}_H + T\lambda_L/\lambda_H ;$$

2. Type L consumption is feasible given T :

$$\pi_{Lb}c_b + \pi_{Lg}c_g \leq \bar{e}_L - T ;$$

3. Type H individuals prefer their own contract to the type L contract:

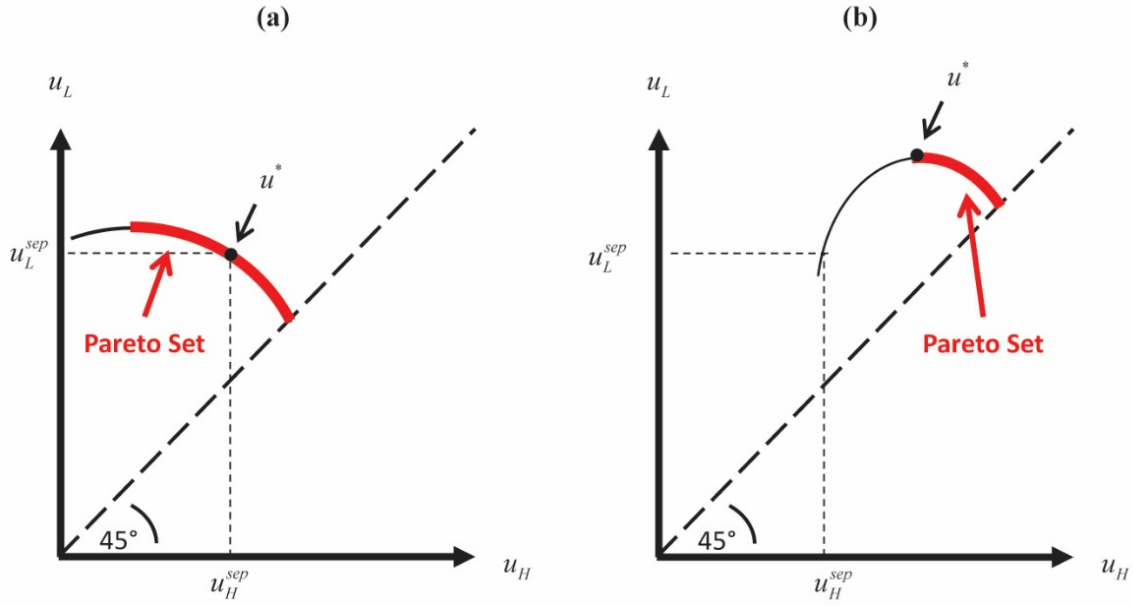
$$U_H(c_H, c_H) \geq U_H(c_b, c_g) ;$$

4. Type L individuals prefer their own contract to the type H contract:

$$U_L(c_b, c_g) \geq U_L(c_H, c_H).$$

To find the Pareto optimum utility sets, the utility of type L is maximized given the utility level of type H , which is pinned down by T . This function is denoted by $\hat{u}_L(u_H)$. The nature of this function depends on the fraction of low-risk types, λ_L . If this fraction is sufficiently small, the curve is as depicted in Figure 3.1(a). The equilibrium utility vector is denoted by u^* . At this point, no transfers between types take place. The Pareto set, depicted in red in Figure 3.1, is the decreasing portion of this concave function.

Figure 3.1: $\hat{u}_L(u_H)$ in the Two-Type Rothschild and Stiglitz Environment for (a) Small λ_L ; and (b) Large λ_L .



The greater the population fraction of the low-risk people, the higher the intersection with the 45 degree line, where all consume the mean endowment independent of type and individual endowment realization. For a sufficiently large λ_L , the Pareto set is a point set with everyone getting utility $v(\bar{e}(\lambda_L))$ with certainty.

3.4.2 Relation to the Rothschild and Stiglitz Equilibrium

Only for the fractions λ_L for which the slope of the curve is negative at u^{sep} as in Figure 3.1(a), does the Rothschild and Stiglitz equilibrium exist. In these cases, the Rothschild and Stiglitz equilibrium and ours are the same. But our equilibrium always exists and is unique.

3.4.3 The General Rothschild and Stiglitz Case

The equilibrium in the more general case can be found by solving the problem of minimizing the l_1 norm of the transfer vector over the set of Pareto optima. The minimum exists given that it is a convex minimization problem with a compact constraint set and a continuous objective function. In some cases, the optimum will have non-degenerate lotteries conditional on the endowment, namely when different degrees of risk aversion can be exploited to separate the types.

3.5 Application to the Spence Signaling Environment

We turn now to the simple [Spence \(1973\)](#) signaling equilibrium, which does not require lotteries to convexify the economy. A point in the commodity space is $x = (c, s)$ where c is consumption and s is the signal and the commodity space is \mathbb{R}^2 .

Types are denoted by $i \in I = \{1, 2\}$. The consumption set is $X = \{x \in \mathbb{R}_+^2 : x \leq x^{max}\}$. Individuals of type 2 are the high-productivity individuals (i.e. $\pi_2 > \pi_1 > 0$) and have the lower disutility of signaling ($0 < \theta_2 < \theta_1$). The utility functions are

$$u_i = c - \theta_i s \quad \forall i \in I.$$

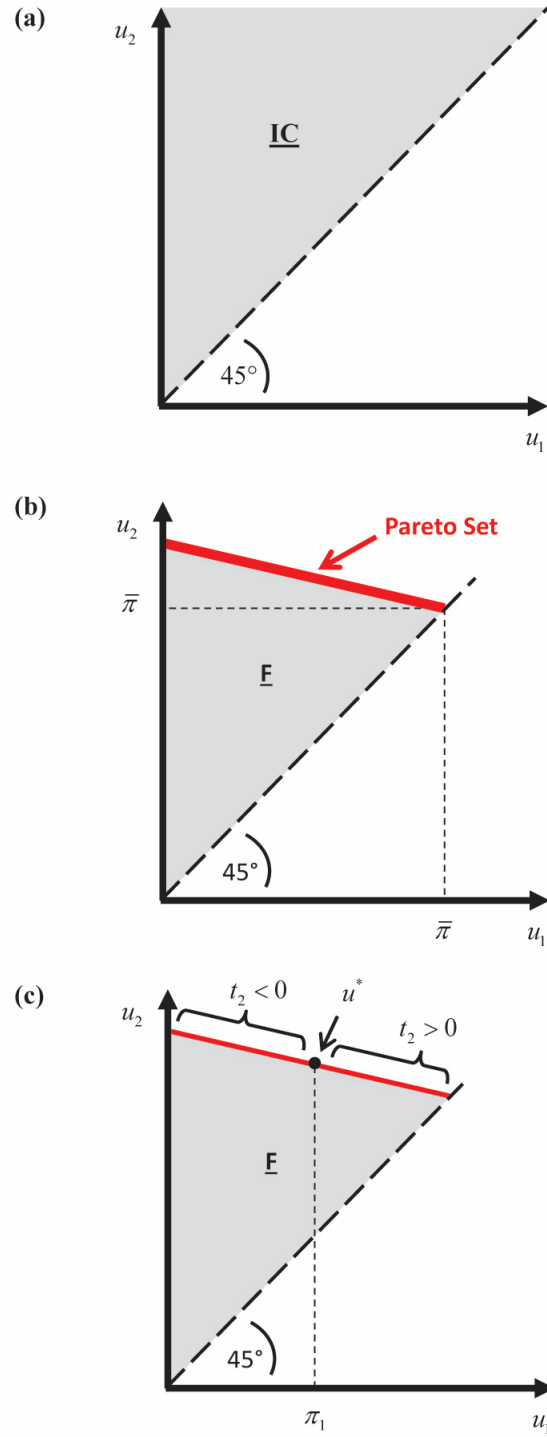
The x^{max} is sufficiently large that all resource feasible allocations are in this consumption set. A sufficiently large c^{max} is π_2 . A sufficiently large s^{max} is π_2/θ_2 . Aggregate transfers from the high-productivity type to the low-productivity type are $t_2(u) = \pi_2 - c_2(u)$.

The equilibrium utility vector is u^* . Figure 3.2(a) specifies the set of utility allocations that are incentive feasible. The set is a convex cone. Figure 3.2(b) specifies the feasible set of allocations and (in red) the Pareto set. For this example, the line

defined by the resource constraint being binding has a negative slope, which requires the fraction of high-productivity types to be below a critical value. If it is not below this critical value, the line has a weakly positive slope, and the Pareto set is a point set with everyone consuming the population average productivity and having a zero signal.

The interesting case is the one in which the Pareto set is a downward-sloping line. Figure 3.2(c) specifies the sign of the transfers from the high-productivity to low-productivity people along the Pareto set. The utility vector for the equilibrium is the point u^* .

Figure 3.2: The Utility Space in the Two-Type Spence Signaling Environment
 Note: Displays (a) The Incentive Compatibility, IC , set; (b) The Pareto set, P ; (c) Transfers along the Pareto set.



3.5.1 *Relation to Spence's Valuation Equilibrium*

Spence used a valuation equilibrium with externalities. The equilibrium depends on the commodity space. The commodities were the set of signals permitted, S , and the consumption good. Let c be the numeraire and w_s the wage for labor of type s . A Spence equilibrium requires that all individuals select signals that maximize their utility given the prices and, for every s that was chosen by some type, the average productivity of those who choose that s is equal to w_s .

Any feasible allocation is an equilibrium for an appropriately selected S as was shown in Prescott and Townsend (1981). For each signal, the associated wage is the average productivity of people who choose that signal.

3.6 Conclusion

The equilibrium concept developed requires there be a single resource constraint and that individuals maximize expected utility. This is an important class of environments that have received a great deal of attention in the economics literature. Equilibrium concepts that are useful in some environments are not useful in others. We see this equilibrium as expanding the set of environments for which there is a useful equilibrium concept.

REFERENCES

- Akerlof, G. A., “The market for “lemons”: Quality uncertainty and the market mechanism”, *The Quarterly Journal of Economics* **84**, 3, 488–500, URL <http://qje.oxfordjournals.org/content/84/3/488.abstract> (1970).
- Badel, A. and M. Huggett, “Taxing top earners: a human capital perspective”, *Federal Reserve Bank of St. Louis Working Paper*, 2014-017, 2014–017 (2014).
- Ben-Porath, Y., “The production of human capital and the life cycle of earnings”, *The Journal of Political Economy* pp. 352–365 (1967).
- Blandin, A., “Learning by doing and ben-porath: Life-cycle predictions and policy analysis”, (2015).
- Boyd, J. H., E. C. Prescott and B. D. Smith, “Organizations in economic analysis”, *Canadian Journal of Economics* pp. 477–491 (1988).
- Brüggemann, B. and J. Yoo, “Aggregate and distributional effects of increasing taxes on top income earners”, (2015).
- Chang, Y., J. F. Gomes and F. Schorfheide, “Learning-by-doing as a propagation mechanism”, *The American Economic Review* **92**, 5, 1498–1520 (2002).
- Cornet, B., “General equilibrium theory and increasing returns: Presentation”, *Journal of Mathematical Economics* **17**, 2-3, 103–118 (1988).
- Debreu, G., “Valuation equilibrium and pareto optimum”, *Proceedings of the national academy of sciences* **40**, 7, 588–592 (1954).
- Diamond, P. A. and E. Saez, “The case for a progressive tax: from basic research to policy recommendations”, *Journal of Economic Perspectives* **25**, 165–190 (2011).
- Edgeworth, F. Y., *Mathematical psychics: An essay on the application of mathematics to the moral sciences*, vol. 10 (Kegan Paul, 1881).
- Ehrenfreund, M. and A. Gearan, “Clinton is open to raising social security taxes on six-figure earners”, *The Washington Post* URL <https://www.washingtonpost.com/news/wonk/wp/2015/08/12/clinton-is-open-to-raising-social-security-taxes-on-six-figure-earners/>, visited on 03/30/2016. (2015).
- Fan, X., A. Seshadri and C. Taber, “Estimation of a life-cycle model with human capital, labor supply and retirement”, (2015).
- Gemici, A. and M. Wiswall, “Evolution of gender differences in post-secondary human capital investments: College majors”, *International Economic Review* **55**, 1, 23–56 (2014).

- Guner, N., R. Kaygusuz and G. Ventura, “Income taxation of us households: Facts and parametric estimates”, *Review of Economic Dynamics* **17**, 4, 559–581 (2014).
- Guner, N., M. Lopez-Daneri and G. Ventura, “Heterogeneity and government revenues: Higher taxes at the top?”, USC-INET Research Paper , 15-23 (2015).
- Guvenen, F., B. Kuruscu and S. Ozkan, “Taxation of human capital and wage inequality: A cross-country analysis”, *The Review of Economic Studies* **81**, 2, 818–850 (2014).
- Hansen, G. D. and S. Imrohoroglu, “Business cycle fluctuations and the life cycle: How important is on-the-job skill accumulation?”, *Journal of Economic Theory* **144**, 6, 2293–2309 (2009).
- Heckman, J., L. Lochner and R. Cossa, “Learning-by-doing vs. on-the-job training: Using variation induced by the eite to distinguish between models of skill formation”, Tech. rep., National Bureau of Economic Research (2002).
- Heckman, J. J., “A life-cycle model of earnings, learning, and consumption”, *Journal of political economy* **84**, 4, S9–S44 (1976).
- Heckman, J. J., L. Lochner and C. Taber, “Tax policy and human capital formation”, Tech. rep., National Bureau of Economic Research (1998).
- Hong, J. H. and J.-V. Ríos-Rull, “Social security, life insurance and annuities for families”, *Journal of Monetary Economics* **54**, 1, 118–140 (2007).
- Huggett, M., G. Ventura and A. Yaron, “Human capital and earnings distribution dynamics”, *Journal of Monetary Economics* **53**, 2, 265–290 (2006).
- Huggett, M., G. Ventura and A. Yaron, “Sources of lifetime inequality”, *American Economic Review* **101**, 7, 2923–54 (2011).
- Imai, S. and M. P. Keane, “Intertemporal labor supply and human capital accumulation*”, *International Economic Review* **45**, 2, 601–641 (2004).
- Killingsworth, M. R., “Learning by doing and investment in training: A synthesis of two rival models of the life cycle”, *The Review of Economic Studies* **49**, 2, 263–271 (1982).
- Kindermann, F., “Earnings related pension schemes and human capital formation”, *Journal of Pension Economics and Finance* **14**, 01, 19–54 (2015).
- Kindermann, F. and D. Krueger, “High marginal tax rates on the top 1%? lessons from a life cycle model with idiosyncratic income risk”, Tech. rep., National Bureau of Economic Research (2014).
- Krueger, D. and A. Ludwig, “Optimal progressive labor income taxation and education subsidies when education decisions and intergenerational transfers are endogenous”, *The American Economic Review* **103**, 3, 496–501 (2013).

- OECD, “Pensions at a glance 2013: Oecd and g20 indicators”, Tech. rep., URL http://dx.doi.org/10.1787/pension_glance-2013-en. (2013).
- Peterman, W., “The effect of endogenous human capital accumulation on optimal taxation”, (2012).
- Prescott, E. C. and R. M. Townsend, “Pareto optima and competitive equilibria with adverse selection and moral hazard”, *Econometrica: Journal of the Econometric Society* pp. 21–45 (1984).
- Rothschild, M. and J. Stiglitz, *Equilibrium in competitive insurance markets: An essay on the economics of imperfect information* (Springer, 1976).
- Shaw, K. L., “Life-cycle labor supply with human capital accumulation”, *International Economic Review* pp. 431–456 (1989).
- Spence, M., “Job market signaling”, *The quarterly journal of Economics* pp. 355–374 (1973).
- Wallenius, J., “Human capital accumulation and the intertemporal elasticity of substitution of labor: How large is the bias?”, *Review of Economic Dynamics* **14**, 4, 577–591 (2011).
- Wallenius, J., “Social security and cross-country differences in hours: A general equilibrium analysis”, *Journal of Economic Dynamics and Control* **37**, 12, 2466–2482 (2013).
- Whitman, K. and D. Shoffner, “The evolution of social security’s taxable maximum”, *Policy Brief*, 2011-02, 1–02 (2011).

APPENDIX A

THE HYBRID HUMAN CAPITAL MODEL

In this paper I analyze the life-cycle patterns for earnings, hourly wages, and hours worked generated by a BP and LBD model, and compare these predictions to data on US workers. An alternative approach to comparing these two models would be to nest each model within a more general human capital function, and then ask what parametrizations of this more general function are consistent with data on US workers. This section lays out such a function and characterizes the optimal decisions of workers when this function is present.

In Section 1.2.2 I laid out the BP version of the human capital function H :

$$H = ah^\phi s^\theta ,$$

and the LBD version of H :

$$H = ah^\phi n^\theta .$$

Recall that a is the learning ability of a worker, and h is the existing human capital of a worker. The two versions of the human capital function accept different inputs: s is human capital investment, and n is time spent producing by the worker. The theoretical distinction between production and investment is that a worker is paid a wage w for the time he spends producing, but is not compensated for time he spends investing.

Each of these versions can be nested by a constant elasticity of substitution (CES) function H which accepts two inputs two human capital formation, production and investment:

$$H(\alpha, \rho) = ah^\phi [\alpha s^\rho + (1 - \alpha)n^\rho]^{\theta/\rho} . \quad (\text{A1})$$

Here, $\alpha \in [0, 1]$ is the weight placed on the investment input relative to the production input, and $\frac{1}{1-\rho} > 0$ is the elasticity of substitution between investment and production. The function in (A1) nests the BP human capital function when $\alpha = 1$

and nests the LBD function when $\alpha = 0$. Killingsworth (1982) discusses some implications of a hybrid human capital function which is a special case of the one in (A1).

Figure A1: $H(\alpha)$ as a Function of Investment

Note: Displays values of newly-produced human capital H as a function of the share of non-leisure time spent investing, $\frac{s}{s+n}$. The total amount of non-leisure time $s + n$ is held fixed. The profiles for different values of α are displayed; all profiles are constructed imposing $\rho \rightarrow 0$.

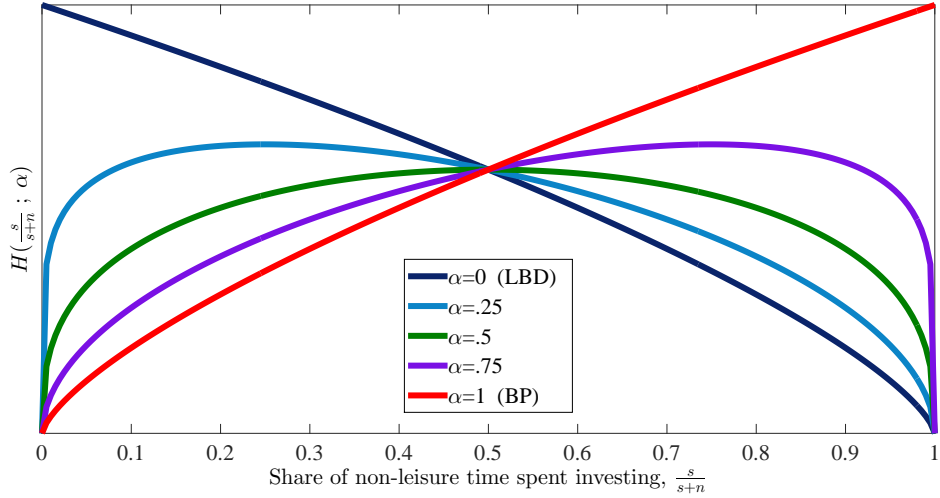


Figure A1 plots the value of H as a function of the share of non-leisure time a worker spends investing, holding the total amount of non-leisure time fixed, and setting $\rho \rightarrow 0$ (i.e. the Cobb-Douglas case). The value of α places a ceiling on the share of non-leisure time a worker will choose to spend investing in a given period. Specifically, given $\alpha \in (0, 1)$, the fraction of time a worker spends investing will always fall in the interval

$$\frac{s}{s+n} \in \left[0, \frac{1}{\left(\frac{1-\alpha}{\alpha}\right)^{\frac{1}{1-\rho}} + 1} \right). \quad (\text{A2})$$

Note that the right hand side of the interval is increasing in α since $\rho < 1$. The reason, as Figure A1 demonstrates for several different values of α , is that H is only increasing in $\frac{s}{s+n}$ over the interval in (A2). Therefore, if a worker chooses to spend

a larger share of non-leisure time than in (A2), he could consume the same amount of leisure, produce the same amount of human capital, and earn strictly more in the current period by spending a larger share of his non-leisure time producing rather than investing. In the limit as $\alpha \rightarrow 0$ (the LBD case) the interval in (A2) collapses to the point 0, since investment is useless for producing human capital in this case.

The following first order conditions pin down the worker's optimal decision:

$$\underbrace{-\frac{\partial u_j}{\partial n_j}}_{\text{MV(Leisure)}} = \underbrace{\lambda \omega h_j}_{\text{MV(Earning)}} + \underbrace{\lambda \left(\frac{\partial h_{j+1}}{\partial n_j} \right) \left[\sum_{k=1}^{J-j} R^{-k} \omega n_{j+k} \left(\frac{\partial h_{j+k}}{\partial h_{j+1}} \right) \right]}_{\text{MV(Learning by producing)}} \quad (\text{A3})$$

$$= \underbrace{\lambda \left(\frac{\partial h_{j+1}}{\partial s_j} \right) \left[\sum_{k=1}^{J-j} R^{-k} \omega n_{j+k} \left(\frac{\partial h_{j+k}}{\partial h_{j+1}} \right) \right]}_{\text{MV(Learning by investing)}} \quad (\text{A4})$$

The above equations resemble the BP first order conditions in the sense that three marginal values are equated: the value of leisure, the value of producing, and the value of investing. As with the BP model, all else equal the marginal value of learning by investing tends to fall as workers age because the number of periods during which new human capital can be used is declining to zero. As a result, workers will typically invest less as they age to satisfy the second equality.

Different from the BP conditions, but similar to the LBD conditions, the marginal value of producing is composed of two terms: the marginal value of earning and the marginal value of learning by producing. Therefore, as with the LBD model the value of producing depends not only on the worker's hourly wage rate ωh_j , but also on the value of learning via producing.

APPENDIX B

VARIANCE OF EARNINGS, WAGES, AND HOURS WORKED

Figure B1: Life-cycle variance of log Earnings, Wages, and Hours in the Data
 Note: Displays the age effects for the variance of log earnings, wages, and hours in the data after controlling for year effects.



APPENDIX C

ROBUSTNESS FOR THE LIFE-CYCLE VARIANCE OF EARNINGS GROWTH

Figure C1: Life-cycle Variance of 5-year Growth Rates

Note: Displays the age effects for the cross-sectional variance of earnings growth in the data after controlling for year effects. Restrictions on hours, wages, and earnings are identical to those listed in Section 4, except that the minimum hours requirement is raised from 500 to 1,800 annual hours. The number of valid observations is 38,464, down from 42,667 for the data corresponding to Figure 4.

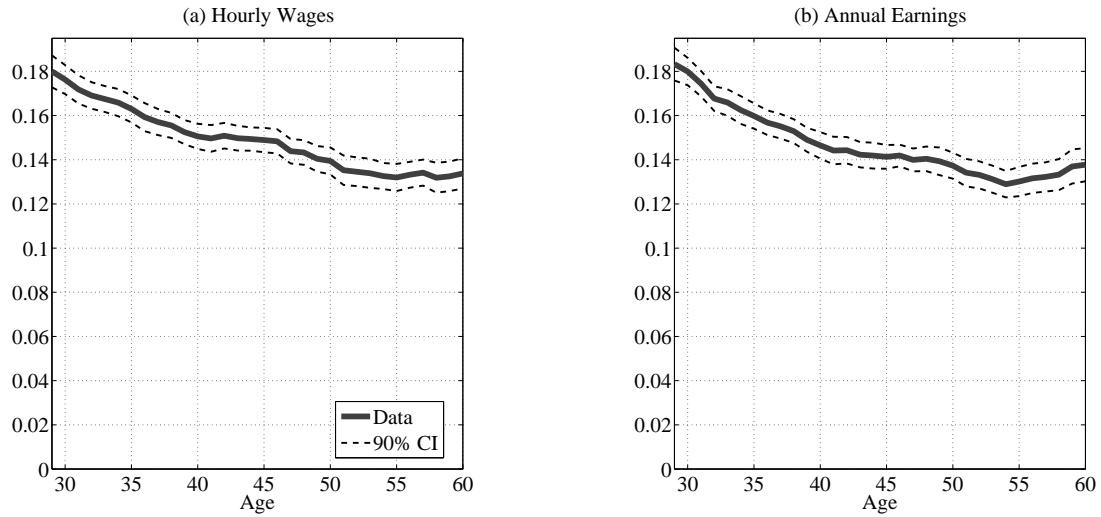


Figure C2: Variance of 5-year Growth Rates by Age, Education

Note: Displays the age effects for the cross-sectional variance of earnings growth in the data after controlling for year effects. Restrictions on hours, wages, and earnings are identical to those listed in Section 4. Education is measured as the number of grades completed (see the variable series which has label 30010 for year 1968). The “No HS” category includes individuals who have completed 0-11 grades; the “HS” category includes individuals who have completed 12-13 grades; the “College” category includes individuals who have completed more than 13 grades. The number of valid observations is 4,269 for the “No HS” category, 16,389 for the “HS” category, and 22,009 for the “College” category.

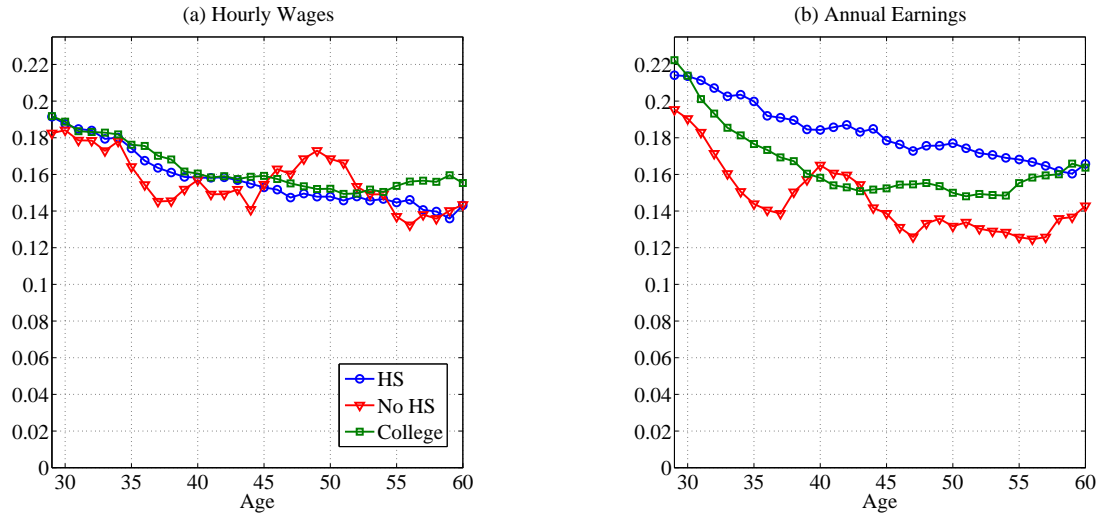


Figure C3: Variance of 5-year Growth Rates by Experience, Education

Note: Displays the age effects for the cross-sectional variance of earnings growth in the data after controlling for year effects. Restrictions on hours, wages, earnings, and education are equivalent to those in Figure A3. Experience is measured as potential experience: age minus grades completed minus 6. The number of valid observations is 5,954 for the “No HS” category, 25,422 for the “HS” category, and 30,518 for the “College” category.

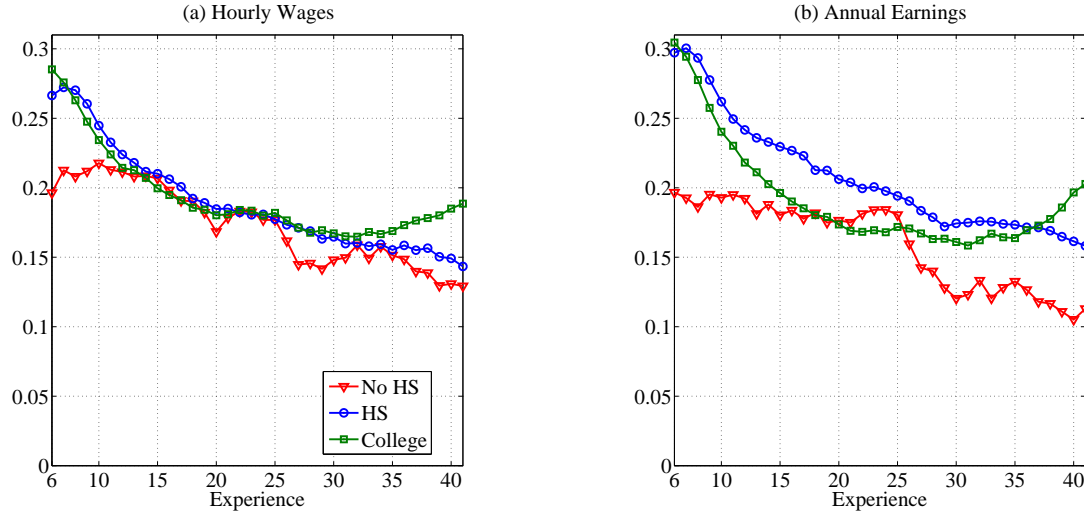


Figure C4: Life-cycle Variance of 5-year Growth Rates, $w_j/w_{j-5} \geq .9$

Note: Displays the age effects for the cross-sectional variance of earnings growth in the data after controlling for year effects. Restrictions on hours, wages, and earnings are identical to those listed in Section 4, except that the minimum growth cutoff is raised from .33 to .9. The number of valid observations is 33,803, down from 42,667 for the data corresponding to Figure 4.

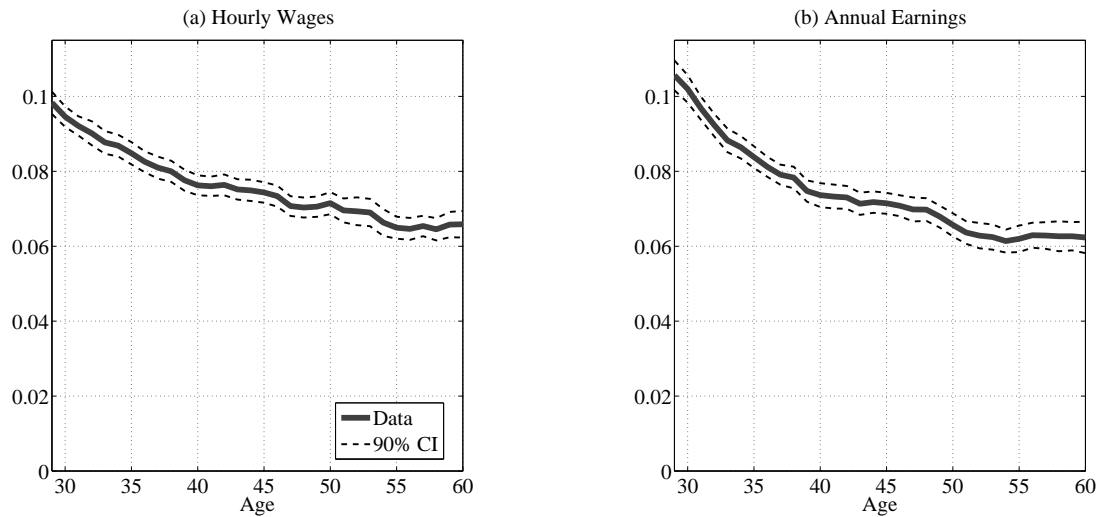
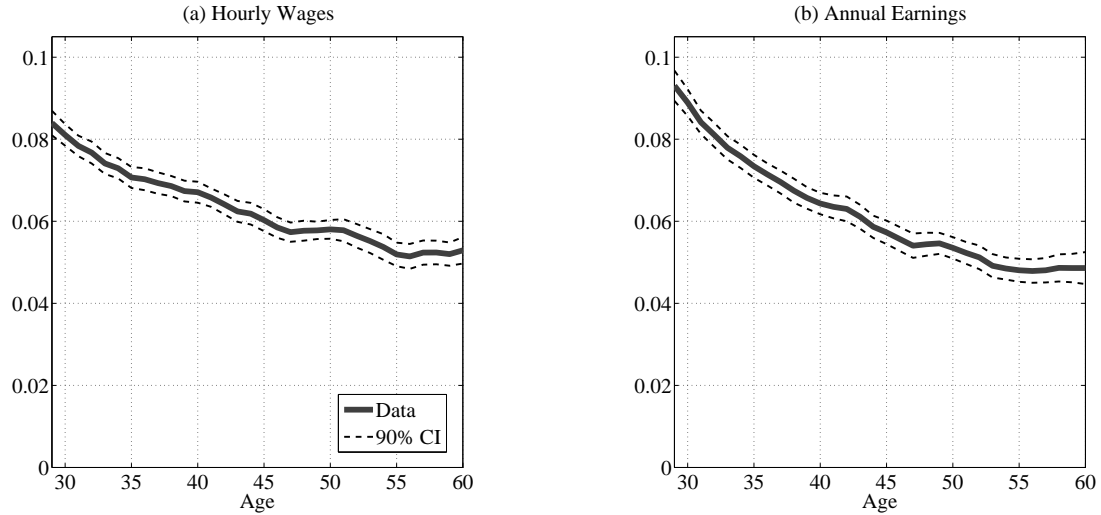


Figure C5: Life-cycle Variance of 5-year Smoothed Growth Rates, $w_j/w_{j-5} \geq .9$

Note: Displays the age effects for the cross-sectional variance of earnings growth in the data after controlling for year effects. Hourly wages and earnings are smoothed with a 3 year centered MA process. Restrictions on hours, smoothed wages, and smoothed earnings are identical to those listen in Section 4, except that the cutoff for wage and earnings growth is raised from .33 to .9. The number of valid observations is 36,103, down from 42,667 for the data corresponding to Figure 4.



This section details several robustness checks for the variance of wage and earnings growth in Figure 1.3.

First, Figure C1 shows that there is a modest change in the level of the profiles, but no substantive change in the life-cycle decline of the profiles, when the minimum hours threshold is increased from 500 hours to 1,800 hours (equivalent to working 35 hours per week for 50 weeks). Therefore, the relatively high variance in growth rates early in the life-cycle is not due to some young workers earning higher wages as they make the transition from part time to full time employment, a phenomenon which does not occur in the human capital model.

Some of the variance in wage growth could be due to differences in the labor markets faced by different education groups. Since the model has only one type of human capital and one labor market, it cannot address heterogeneity in wage growth which arises because of differences in labor markets. However, Figure C2 reveals that when I recompute the statistic within different education groups the initial level and life-cycle decline of the profiles remain broadly consistent with their values in Figure 1.3.¹ Additionally, Figure C3 demonstrates that the same qualitative pattern holds when I estimate potential experience effects within education groups instead of age effects.

Another possibility is that the high variance in wage growth for young adults could be due to more frequent job losses or changes in occupations relative to older adults, which often entail substantial declines in wages and earnings in the short run. To remove the component of growth variance which is due to large wage declines, I throw out all observations in which wages or earnings fell by more than 10%, and plot the resulting age effects in Figure C4. The initial levels of the profiles are substantially lower than the initial levels in Figure 1.3, but the size of the declines over the life-cycle are similar: the variance in hourly wage growth falls by .034 in Figure C4(a) versus .046 in Figure 1.3(a); the variance in earnings growth falls by .044 in Figure C4(b) versus .063 in Figure 1.3(b). This establishes that the decline in the variance of wage and earnings growth is not driven by a reduction in the frequency of large declines in wages or earnings.

¹The profile for individuals without a high school degree is somewhat erratic due to a low sample size.

A final concern is that the life-cycle decline in these profiles is due to a decline in the variance of temporary idiosyncratic shocks to earnings, which are not present in my (deterministic) model. This suggests it may be important to distinguish between temporary wage growth and persistent wage growth. To remove purely temporary changes I smooth the wages and earnings of all individuals with a 3 year Moving Average process, then recompute the growth rates using these smoothed series. I also exclude observations in which earnings or wages decline by more than 10%, as was discussed in the preceding paragraph. The age effects obtained from equation for the variance in smoothed growth rates are plotted in Figure C5. The initial level of the profiles are roughly half the initial levels from Figure 1.3, but the life-cycle decline in the profiles are similar. The profile for hourly wages declines by .033 in Figure C5(a) compared to .046 in Figure 1.3(a); the profile for annual earnings declines by .045 in Figure C5(b) versus .063 in Figure 1.3(b).